Symmetry and Particle Physics

Michaelmas Term 2007

Jan B. Gutowski
DAMTP, Centre for Mathematical Sciences
University of Cambridge
Wilberforce Road, Cambridge
CB3 0WA, UK
Email: J.B.Gutowski@damtp.cam.ac.uk

Contents

Recommended Books

General Particle Physics Books

- ⋆ D. H. Perkins, *Introduction to High energy Physics,* 4th ed., CUP (2000).
- • B. R. Martin and G. Shaw, *Particle Physics,* 2nd ed., Wiley (1998).

Lie Algebra books written for Physicists

- ⋆ H. Georgi, *Lie Algebras in Particle Physics,* Perseus Books (1999).
- ⋆ J. Fuchs and C. Schweigert, *Symmetries, Lie Algebras and Representations,* 2nd ed., CUP (2003).
- • H. F. Jones, *Groups, Representations and Physics,* 2nd ed., IOP Publishing (1998).
- • J. Cornwell, *Group Theory in Physics,* (Volume 2), Academic Press (1986).

Pure Mathematics Lie Algebra books

- • S. Helgason, *Differential geometry, Lie groups, and symmetric spaces,* 3rd ed., Academic Press (1978).
- ⋆ H. Samelson, *Notes on Lie Algebras,* Springer (1990).
- • W. Fulton and J. Harris, *Representation Theory, A First Course,* 3rd ed. Springer (1991).

For an introduction to some aspects of Lie group differential geometry not covered in this course:

- • M. Nakahara, *Geometry, Topology and Physics,* 2nd ed., Institute of Physics Publishing (2003).

References for Spacetime Symmetry and Gauge Theory Applications

- • T-P. Cheng and L-F. Li, *Gauge Theory of Elementary Particle Physics,* Oxford (1984).
- ⋆ S. Pokorski, *Gauge Field Theories,* 2nd ed., CUP (2000).
- • S. Weinberg, *The Quantum Theory of Fields,* (Book 1), CUP (2005).
- ⋆ J. Buchbinder and S. Kuzenko, *Ideas and Methods of Supersymmetry and Supergravity, Or a Walk Through Superspace,* 2nd ed., Institute of Physics Publishing (1998).

1. Introduction to Symmetry and Particles

Symmetry simplifies the description of physical phenomena. It plays a particularly important role in particle physics, for without it there would be no clear understanding of the relationships between particles. Historically, there has been an "explosion" in the number of particles discovered in high energy experiments since the discovery that atoms are not fundamental particles. Collisions in modern accelerators can produce cascades involving hundreds of types of different particles: $p, n, \Pi, K, \Lambda, \Sigma \ldots$ etc.

The key mathematical framework for symmetry is group theory: symmetry transformations form groups under composition. Although the symmetries of a physical system are not sufficient to fully describe its behaviour - for that one requires a complete dynamical theory - it is possible to use symmetry to find useful constraints. For the physical systems which we shall consider, these groups are smooth in the sense that their elements depend smoothly on a finite number of parameters (called co-ordinates). These groups are Lie groups, whose properties we will investigate in greater detail in the following lectures. We will see that the important information needed to describe the properties of Lie groups is encoded in "infinitessimal transformations", which are close in some sense to the identity transformation. The properties of these transformations, which are elements of the tangent space of the Lie group, can be investigated using (relatively) straightforward linear algebra. This simplifies the analysis considerably. We will make these rather vague statements more precise in the next chapter.

Examples of symmetries include

i) Spacetime symmetries: these are described by the Poincaré group. This is only an approximate symmetry, because it is broken in the presence of gravity. Gravity is the weakest of all the interactions involving particles, and we will not consider it here.

ii) Internal symmetries of particles. These relate processes involving different types of particles. For example, isospin relates u and d quarks. Conservation laws can be found for particular types of interaction which constrain the possible outcomes. These symmetries are also approximate; isospin is not exact because there is a (small) mass difference between m_u and m_d. Electromagnetic effects also break the symmetry.

iii) Gauge symmetries. These lead to specific types of dynamical theories describing types of particles, and give rise to conserved charges. Gauge symmetries if present, appear to be exact.

1.1 Elementary and Composite Particles

The fundamental particles are quarks, leptons and gauge particles.

The *quarks* are spin 1/2 fermions, and can be arranged into three families

						Electric Charge (e)
u	(0.3 GeV)	c	(1.6 GeV)	t	(175 GeV)	$\frac{2}{3}$
d	($\approx$ 0.3 GeV)	s	(0.5 GeV)	b	(4.5 GeV)	$-\frac{1}{3}$

The quark labels u, d, s, c, t, b stand for up, down, strange, charmed, top and bottom. The quarks carry a fractional electric charge. Each quark has three *colour* states. Quarks are not seen as free particles, so their masses are ill-defined (the masses above are "effective" masses, deduced from the masses of composite particles containing quarks).

The *leptons* are also spin 1/2 fermions and can be arranged into three families

			Electric Charge (e)
e^- (0.5 MeV)	μ^- (106 MeV)	τ^- (1.8 GeV)	-1
ν_e (< 10 eV)	ν_μ (< 0.16 MeV)	ν_τ (< 18 MeV)	0

The leptons carry integral electric charge. The muon μ and taon τ are heavy unstable versions of the electron e. Each flavour of charged lepton is paired with a neutral particle ν, called a neutrino. The neutrinos are stable, and have a very small mass (which is taken to vanish in the standard model).

All these particles have *antiparticles* with the same mass and opposite electric charge (conventionally, for many particles, the antiparticles carry a bar above the symbol, e.g. the antiparticle of u is $\bar{u}$). The antiparticles of the charged leptons are often denoted by a change of $-$ to $+$, so the positron e^+ is the antiparticle of the electron e^- etc. The antineutrinos $\bar{\nu}$ differ from the neutrinos ν by a change in helicity (to be defined later...).

Hadrons are made from bound states of quarks (which are colour neutral singlets).

i) The *baryons* are formed from bound states of three quarks qqq; antibaryons are formed from bound states of three antiquarks $\bar{q}\bar{q}\bar{q}$

For example, the nucleons are given by

$$\begin{cases} p = uud & : \quad 938\ Mev \\ n = udd & : \quad 940\ Mev \end{cases}$$

ii) *Mesons* are formed from bound states of a quark and an antiquark $q\bar{q}$.

For example, the pions are given by

$$\begin{cases} \pi^+ = u\bar{d} & : \quad 140\ Mev \\ \pi^- = d\bar{u} & : \quad 140\ Mev \\ \pi^0 = u\bar{u}, d\bar{d}\ \text{superposition} & : \quad 135\ Mev \end{cases}$$

Other particles are made from heavy quarks; such as the strange particles $K^+ = u\bar{s}$ with mass 494 Mev , $\Lambda = uds$ with mass 1115 Mev, and Charmonium $\psi = c\bar{c}$ with mass 3.1 Gev.

The *gauge particles* mediate forces between the hadrons and leptons. They are bosons, with integral spin.

	Mass (GeV)	Interaction
γ (photon)	0	Electromagnetic
W^+	80	Weak
W^-	80	Weak
Z^0	91	Weak
g (gluon)	0	Strong

The gluons are responsible for interquark forces which bind quarks together in nucleons. It is conjectured that a spin 2 gauge boson called the graviton is the mediating particle for gravitational forces, though detecting this is extremely difficult, due to the weakness of gravitational forces compared to other interactions.

1.2 Interactions

There are three types of interaction which are of importance in particle physics: the strong, electromagnetic and weak interactions.

1.2.1 The Strong Interaction

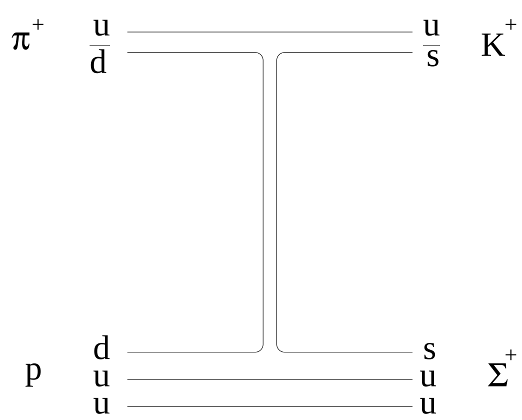
The strong interaction is the strongest interaction.

- Responsible for binding of quarks to form hadrons (electromagnetic effects are much weaker)
- Dominant in scattering processes involving just hadrons. For example, $pp \to pp$ is an elastic process at low energy; whereas $pp \longrightarrow pp\pi^+\pi^-$ is an inelastic process at higher energy.
- Responsible for binding forces between nucleons p and n, and hence for all nuclear structure.

Properties of the Strong Interaction:

i) The strong interaction preserves quark flavours, although $q\bar{q}$ pairs can be produced and destroyed provided q, $\bar{q}$ are the same flavour.

An example of this is:



The Σ^+ and K^+ particles decay, but not via the strong interaction, because of conservation of strange quarks.

ii) Basic strong forces are "flavour blind". For example, the interquark force between $q\bar{q}$ bound states in the $\psi = c\bar{c}$ (charmonium) and $\Upsilon = b\bar{b}$ (bottomonium) mesons are well-approximated by the potential

$$V \sim \frac{\alpha}{r} + \beta r \tag{1.1}$$

and the differences in energy levels for these mesons is approximately the same.

The binding energy differences can be attributed to the mass difference of the b and c quarks.

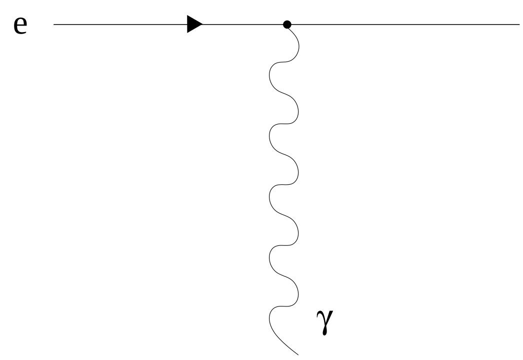
iii) Physics is unchanged if all particles are replaced by antiparticles.

The dynamical theory governing the strong interactions is Quantum Chromodynamics (QCD), which is a gauge theory of quarks and gluons. This is in good agreement with experiment, however non-perturbative calculations are difficult.

1.2.2 Electromagnetic Interactions

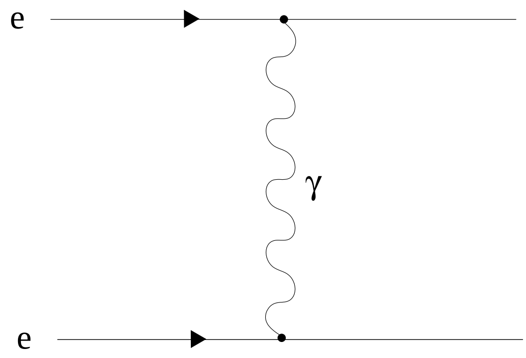
The electromagnetic interactions are weaker than the strong interactions. They occur in the interactions between electrically charged particles, such as charged leptons, mediated by photons.

The simplest electromagnetic process consists of the absorption or emission of a photon by an electron:

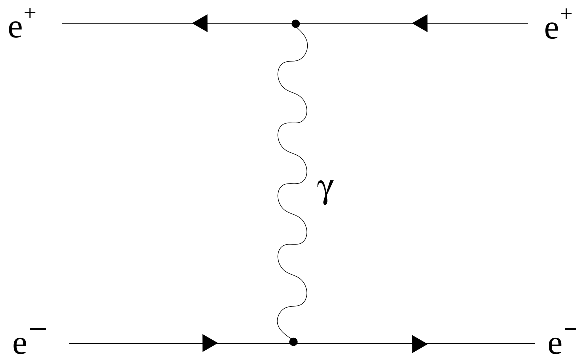


This process cannot occur for a free electron, as it would violate conservation of 4-momentum, rather it involves electrons in atoms, and the 4-momentum of the entire atom and photon are conserved.

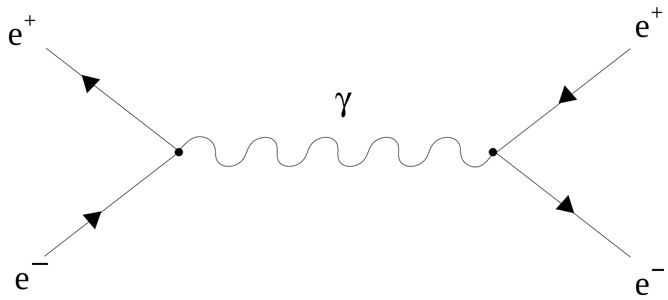
Other examples of electromagnetic interactions are electron scattering mediated by photon exchange



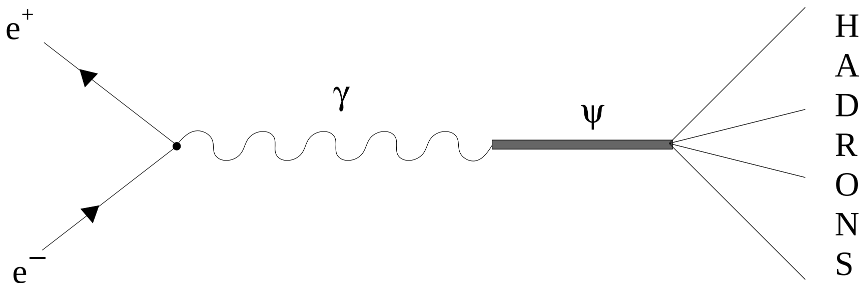
and there are also smaller contributions to this process from multi-photon exchanges. Electron-positron interactions are also mediated by electromagnetic interactions



+



Electron-positron annihilation can also produce particles such as charmonium or bottomonium



The dynamic theory governing electromagnetic interactions is Quantum Electrodynamics (QED), which is very well tested experimentally.

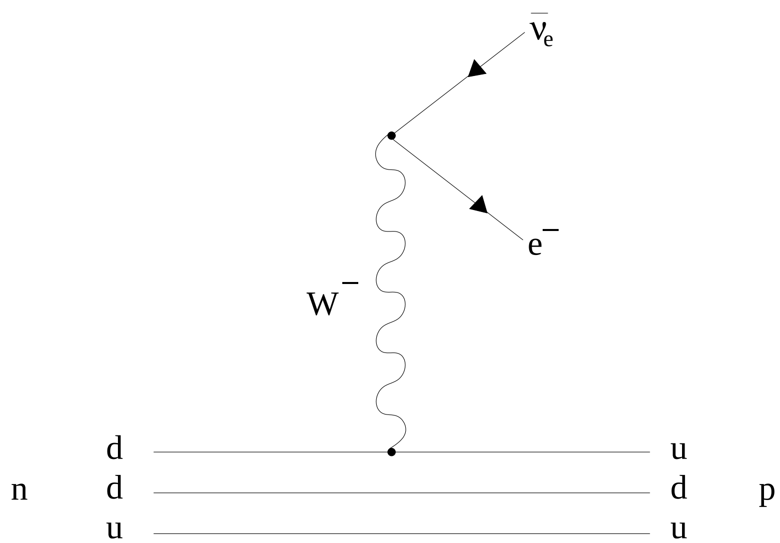
Neutrinos have no electromagnetic or strong interactions.

1.2.3 The weak interaction

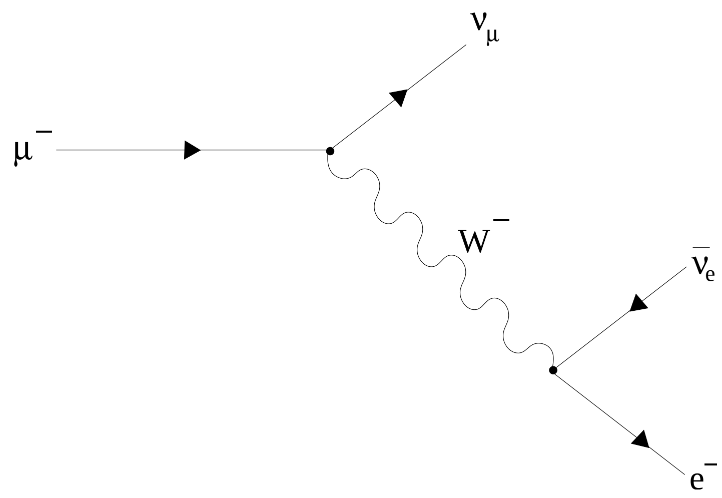
The weak interaction is considerably weaker than both the strong and electromagnetic interactions, they are mediated by the charged and neutral vector bosons $W^{\pm}$ and Z^0 which are very massive and produce only short range interactions. Weak interactions occur between all quarks and leptons, however they are in general negligable when there are strong or electromagnetic interactions present. Only in the absence of strong and electromagnetic interactions is the weak interaction noticable.

Unlike the strong and electromagnetic interactions, weak interactions can involve neutrinos. Weak interactions, unlike strong interactions, can also produce flavour change in quarks and neutrinos.

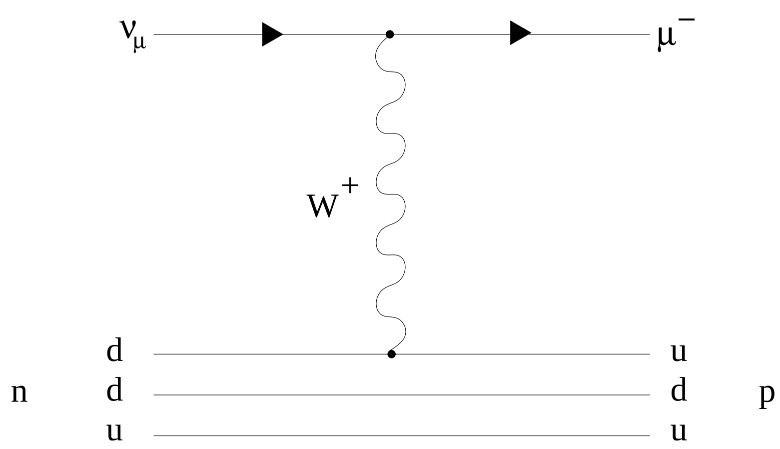
The gauge bosons $W^{\pm}$ carry electric charge and they can change the flavour of quarks. Examples of W-boson mediated weak interactions are $n \longrightarrow p + e^- + \bar{\nu}_e$:



and $\mu^- \longrightarrow e^- + \bar{\nu}_e + \nu_\mu$:



and $\nu_\mu + n \to \mu^- + p$

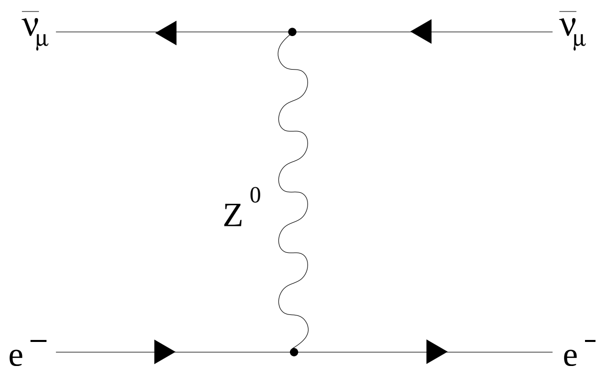


The flavour changes within one family are dominant; e.g.

$$\begin{aligned} e^- &\leftrightarrow \nu_e, \qquad & \mu^- &\leftrightarrow \nu_\mu \\ u &\leftrightarrow d, & c &\leftrightarrow s \end{aligned} \tag{1.2}$$

whereas changes between families, like $u \leftrightarrow s$ and $u \leftrightarrow b$ are "Cabibbo suppressed".

The neutral Z^0, like the photon, does not change quark flavour; though unlike the photon, it couples to neutrinos. An example of a Z^0 mediated scattering process is $\bar{\nu}_\mu e^-$ scattering:



In any process in which a photon is exchanged, it is possible to have a Z^0 boson exchange. At low energies, the electromagnetic interaction dominates; however at high energies and momenta, the electromagnetic and weak interactions become comparable. The unified theory of electromagnetic and weak interactions is Weinberg-Salam theory.

1.2.4 Typical Hadron Lifetimes

Typical hadron lifetimes (valid for most decays) via the three interactions are summarized below:

Interaction	Lifetime (s)
Strong	$10^{-22} - 10^{-24}$
Electromagnetic	$10^{-16} - 10^{-21}$
Weak	$10^{-7} - 10^{-13}$

with the notable exceptional case being weak neutron decay, which has average lifetime of $10^3 s$.

1.3 Conserved Quantum Numbers

Given a configuration of particles containing particle P, we define $N(P)$ to denote the number of P-particles in the configuration. We define various quantum numbers associated with leptons and hadrons.

Definition 1. *There are three lepton numbers. The electron, muon and tauon numbers are given by*

$$\begin{aligned} L_e &= N(e^-) - N(e^+) + N(\nu_e) - N(\bar{\nu}_e) \\ L_\mu &= N(\mu^-) - N(\mu^+) + N(\nu_\mu) - N(\bar{\nu}_\mu) \\ L_\tau &= N(\tau^-) - N(\tau^+) + N(\nu_\tau) - N(\bar{\nu}_\tau) \end{aligned} \tag{1.3}$$

In electromagnetic interactions, where there are no neutrinos involved, conservation of L is equivalent to the statement that leptons and anti-leptons can only be created or annihilated in pairs. For weak interactions there are more possibilities, so for example, an election e^- and anti-neutrino $\bar{\nu}_e$ could be created. Lepton numbers are conserved in all interactions.

There are also various quantum numbers associated with baryons.

Definition 2. *The four quark numbers S, C, $\tilde{B}$ and T corresponding to strangeness, charm, bottom and top are defined by*

$$\begin{aligned} S &= -(N(s) - N(\bar{s})) \\ C &= (N(c) - N(\bar{c})) \\ \tilde{B} &= -(N(b) - N(\bar{b})) \\ T &= (N(t) - N(\bar{t})) \end{aligned} \tag{1.4}$$

These quark quantum numbers, together with $N(u) - N(\bar{u})$ and $N(d) - N(\bar{d})$, are conserved in strong and electromagnetic interactions, because in these interactions quarks and antiquarks are only created or annihilated in pairs. The quark quantum numbers are *not* conserved in weak interactions, because it is possible for quark flavours to change.

Definition 3. *The baryon number B is defined by*

$$B = \frac{1}{3}(N(q) - N(\bar{q})) \tag{1.5}$$

where $N(q)$ and $N(\bar{q})$ are the total number of quarks and antiquarks. Baryons therefore have $B = 1$ and antibaryons have $B = -1$; mesons have $B = 0$. B is conserved in all interactions.

Note that one can write

$$B = \frac{1}{3}(N(u) - N(\bar{u}) + N(d) - N(\bar{d}) + C + T - S - \tilde{B}) \tag{1.6}$$

Definition 4. *The quantum number Q is the total electric charge. Q is conserved in all interactions*

In the absence of charged leptons, such as in strong interaction processes, one can write

$$Q = \frac{2}{3}(N(u) - N(\bar{u}) + C + T) - \frac{1}{3}(N(d) - N(\bar{d}) - S - \tilde{B}) \tag{1.7}$$

Hence, for strong interactions, the four quark quantum numbers S, C, $\tilde{B}$, T together with Q and B are sufficient to determine $N(u) - N(\bar{u})$ and $N(d) - N(\bar{d})$.

2. Elementary Theory of Lie Groups and Lie Algebras

2.1 Differentiable Manifolds

Definition 5. *A n-dimensional real smooth manifold M is a (Hausdorff topological) space which is equipped with a set of open sets U^α such that*

1) For each $p \in M$, there is some U^α with $p \in U^\alpha$

2) For each U^α, there is an invertible homeomorphism $x_\alpha : U^\alpha \to \mathbb{R}^n$ onto an open subset of $\mathbb{R}^n$ such that if $U^\alpha \cap U^\beta \neq \emptyset$ then the map

$$x_\beta \circ {x_\alpha}^{-1} : x_\alpha(U^\alpha \cap U^\beta) \to x_\beta(U^\alpha \cap U^\beta) \tag{2.1}$$

is smooth (infinitely differentiable) as a function on $\mathbb{R}^n$.

The open sets U^α together with the maps x_α are called charts, the set of all charts is called an atlas. The maps x_α are local co-ordinates on M defined on the U^α, and have components x^i_α for $i = 1, \ldots, n$. So a smooth manifold looks locally like a portion of $\mathbb{R}^n$.

A n-dimensional complex manifold is defined in an exactly analogous manner to a real manifold, with $\mathbb{R}^n$ replaced by $\mathbb{C}^n$ throughout.

Definition 6. *Suppose M is a m-dimensional smooth manifold, and N is a n-dimensional smooth manifold, with charts (U^α, x_α), (W^A, y_A) respectively. Then the Cartesian product $X = M \times N$ is a $m+n$-dimensional smooth manifold, equipped with the standard Cartesian product topology.*

The charts are $V^{\alpha,A} = U^\alpha \times W^A$ with corresponding local co-ordinates

$$z_{\alpha,A} = x_\alpha \times y_A : U^\alpha \times W^A \to \mathbb{R}^{m+n} \tag{2.2}$$

Definition 7. *Suppose M is a m-dimensional smooth manifold, and N is a n-dimensional smooth manifold, with charts (U^α, x_α), (W^A, y_A) respectively. Then a function $f : M \to N$ is smooth if for every U^α and W^A such that $f(U^\alpha) \cap W^A \neq \emptyset$, the map*

$$y_A \circ f \circ x_\alpha^{-1} : x_\alpha(U^\alpha) \to y_A(W^A) \tag{2.3}$$

is smooth as a function $\mathbb{R}^m \to \mathbb{R}^n$.

Definition 8. *A smooth curve on a manifold M is a map $\gamma : (a, b) \to M$ where (a, b) is some open interval in $\mathbb{R}$ such that if U is a chart with local co-ordinates x then the map*

$$x \circ \gamma : (a, b) \to \mathbb{R}^n \tag{2.4}$$

may be differentiated arbitrarily often.

2.2 Lie Groups

Definition 9. *A group G is a set equipped with a map $\bullet : G \times G \to G$, called group multiplication, given by $(g_1, g_2) \to g_1 \bullet g_2 \in G$ for g_1, $g_2 \in G$. Group multiplication satisfies* *i) There exists $e \in G$ such that $g \bullet e = e \bullet g = g$ for all $g \in G$. e is called an identity element.*

ii) For every $g \in G$ there exists an inverse $g^{-1} \in G$ such that $g \bullet g^{-1} = g^{-1} \bullet g = e$.

iii) For all $g_1, g_2, g_3 \in G$; $g_1 \bullet (g_2 \bullet g_3) = (g_1 \bullet g_2) \bullet g_3$, so group multiplication is associative.

It is elementary to see that the identity e is unique, and g has a unique inverse g^{-1}.

Definition 10. *A Lie group G is a smooth differentiable manifold which is also a group, where the group multiplication $\bullet$ has the following properties*

i) The map $\bullet : G \times G \to G$ given by $(g_1, g_2) \to g_1 \bullet g_2$ is a smooth map.

ii) The inverse map $G \to G$ given by $g \to g^{-1}$ is a smooth map

Henceforth, we shall drop the $\bullet$ for group multiplication and just write $g_1 \bullet g_2 = g_1 g_2$.
Examples:

Many of the most physically interesting Lie groups are matrix Lie groups in various dimensions. These are subgroups of $GL(n, \mathbb{R})$ (or $GL(n, \mathbb{C})$), the $n \times n$ real (or complex) invertible matrices. Group multiplication and inversion are standard matrix multiplication and inversion.

Suppose that G is a matrix Lie group of dimension k. Let the local co-ordinates be x^i for $i = 1, \ldots, k$. Then $g \in G$ is described by its matrix components $g^{AB}(x^i)$ for $A, B = 1, \ldots, n$. The g^{AB} are smooth functions of the co-ordinates x^i. Examples of matrix Lie groups are (here $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$):

i) $GL(n, \mathbb{F})$, the invertible $n \times n$ matrices over $\mathbb{F}$. The co-ordinates of $GL(n, \mathbb{F})$ are the n^2 real (or complex) components of the matrices.

ii) $SL(n, \mathbb{F}) = \{M \in GL(n, \mathbb{F}) : \det M = 1\}$

iii) $O(n) = \{M \in GL(n, \mathbb{R}) : MM^T = \mathbb{I}_n\}$

iv) $U(n) = \{M \in GL(n, \mathbb{C}) : MM^\dagger = \mathbb{I}_n\}$, where $\dagger$ is the hermitian transpose.

v) $SO(n) = \{M \in GL(n, \mathbb{R}) : MM^T = \mathbb{I}_n \text{ and } \det M = 1\}$

vi) $SU(n) = \{M \in GL(n, \mathbb{C}) : MM^\dagger = \mathbb{I}_n \text{ and } \det M = 1\}$. $SU(2)$ and $SU(3)$ play a particularly important role in the standard model of particle physics.

vii) $SO(1, n-1) = \{M \in GL(n, \mathbb{R}) : M^T \eta M = \eta \text{ and } \det M = 1\}$
where $\eta = \text{diag } (1, -1, -1, \cdots - 1)$ is the n-dimensional Minkowski metric.

There are other examples, some of which we will examine in more detail later. It can be shown that any closed subgroup H of $GL(n, \mathbb{F})$ (i.e. any subgroup which contains all its accumulation points) is a Lie group.

Some of these groups are related to each other by group isomorphism; a particularly simple example is $SO(2) \cong U(1)$. Elements of $U(1)$ consist of unit-modulus complex numbers $e^{i\theta}$ for $\theta \in \mathbb{R}$ under multiplication, whereas $SO(2)$ consists of matrices

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{2.5}$$

which satisfy $R(\theta+\phi) = R(\theta)R(\phi)$. The map $T : U(1) \to SO(2)$ given by $T(e^{i\theta}) = R(\theta)$ is a group isomorphism.

2.3 Compact and Connected Lie Groups

A lie group G is compact if G is compact as a manifold. Recall that a subset of $U \subset \mathbb{R}^n$ is compact iff it is closed and bounded, or equivalently iff every sequence $u_n \in U$ has a subsequence which converges to some $u \in U$.

It is straightforward to see that $SU(n)$ is compact, for if we denote the rows of $M \in SU(n)$ by R_A then $R_A^\dagger R_B = \delta_{AB}$. Hence the components M^{AB} are all bounded $|M^{AB}| \leq 1$. So it follows that if M_n is a sequence of points in $SU(n)$, then by repeated application of the Bolzano-Weierstrass theorem, there is a subsequence M_{n_r} which converges to some matrix N. Moreover as the constraints $\det M_{n_r} = 1$ and $M_{n_r} M_{n_r}^\dagger = 1$ are smooth functions of the matrix components, one must also have $\det N = 1$ and $NN^\dagger = 1$ in the limit as $r \to \infty$, i.e. $N \in SU(n)$.

[There is a subtlety concerning convergence which we have glossed over, namely how one actually defines convergence. We assume the existence of some matrix norm (for example $||M||_{\text{sup}} = \max(|M^{AB}|)$) with respect to which convergence is defined. As all (finite-dimensional) matrix norms are equivalent, convergence with respect to one matrix norm ensures convergence with respect to any norm].

In contrast, the Lorentz group $SO(1, n-1)$ is not compact. For example, consider for simplicity $SO(1, 1)$. One can define a sequence of elements $M_n \in SO(1, 1)$ by

$$M_n = \begin{pmatrix} \cosh n & \sinh n \\ \sinh n & \cosh n \end{pmatrix} \tag{2.6}$$

As the components of M_n are unbounded, it follows that M_n cannot have a convergent subsequence. Observe that as $SO(1, n-1)$ is a Lie subgroup of both $SL(n, \mathbb{R})$ and $GL(n, \mathbb{R})$ it must follow that $SL(n, \mathbb{R})$ and $GL(n, \mathbb{R})$ are also non-compact.

A Lie group G is said to be connected if any two points in the group can be linked together by a continuous curve in G.

$O(n)$ is not connected. To see this, observe that if $M \in O(n)$ then $MM^T = 1$ and on taking the determinant this implies $\det M = \pm 1$. Now take $M \in O(n)$ with $\det M = -1$, so if $O(n)$ is connected, there is a continuous curve $\gamma : [0, 1] \to O(n)$ with $\gamma(0) = \mathbb{I}$ and $\gamma(1) = M$. We can then compute $\det \gamma(t)$ which must be a continuous real function of t such that $\det \gamma(t) \in \{-1, 1\}$ for all $t \in [0, 1]$ and $\det \gamma(0) = 1$, $\det \gamma(1) = -1$. This is not possible.

We shall say that two points in G are connected if they can be linked with a continuous curve. This defines an equivalence relation on G, and hence partitions G into equivalence classes of connected points; the equivalence class of $g \in G$ is called the connected component of g. The equivalence class of points of $O(n)$ connected to $\mathbb{I}$ is $SO(n)$, which is connected.

2.4 Tangent Vectors

Suppose that U is an open subset of a manifold M, and that the curve γ passes through some $p \in U$ with $\gamma(t_0) = p$. Then the curve defines a tangent vector at p, denoted by $\dot{\gamma}_p$, which maps smooth real functions $f : U \to \mathbb{R}$ to $\mathbb{R}$ according to

$$\dot{\gamma}_p : f \to \Big[\frac{d}{dt}(f \circ \gamma(t))\Big]_{t=t_0} \tag{2.7}$$

The components of the tangent vector are

$$\dot{\gamma}_p^m = \Big[\frac{d}{dt}((x \circ \gamma)^m)\Big]_{t=t_0} = \dot{\gamma}_p(x^m) \tag{2.8}$$

Note that one can write (using the chain rule)

$$\begin{aligned} \dot{\gamma}_p(f) &= \Big[\frac{d}{dt}(f \circ \gamma(t))\Big]_{t=t_0} \\ &= \Big[\frac{d}{dt}(f \circ x^{-1} \circ x \circ \gamma(t))\Big]_{t=t_0} \\ &= \sum_{i=1}^{n} \frac{\partial}{\partial x^i}(f \circ x^{-1})|_{x(p)} (\frac{d}{dt}(x \circ \gamma)^i)_{t=t_0} \\ &= \sum_{i=1}^{n} \frac{\partial}{\partial x^i}(f \circ x^{-1})|_{x(p)} \dot{\gamma}_p^i \end{aligned} \tag{2.9}$$

Proposition 1. *The set of all tangent vectors at p forms a n-dimensional vector space (where $n = \dim M$), denoted by $T_p(M)$.*

<u>Proof</u>

Suppose that p lies in the chart U with local co-ordinates x. Suppose also that V, $W \in T_p(M)$ are tangent vectors at p corresponding to the curves γ, σ, where without loss of generality we can take $\gamma : (a, b) \to M$, $\sigma : (a, b) \to M$ with $a < t_0 < b$ and $\gamma(t_0) = \sigma(t_0) = p$

Take $a, b \in \mathbb{R}$. Consider the curve $\hat{\rho}$ in $\mathbb{R}^n$ defined by

$$\hat{\rho}(t) = a(x \circ \gamma)(t) + b(x \circ \sigma)(t) - (a + b - 1)x(p) \tag{2.10}$$

where scalar multiplication and vector addition are the standard operations in $\mathbb{R}^n$. Note that $\hat{\rho}(t_0) = x(p)$.

Then define the curve ρ on U by $\rho = x^{-1} \circ \hat{\rho}$, so that $\rho(t_0) = p$.

If f is a smooth function on U then by (2.9) it follows that

$$\dot{\rho}_p(f) = \sum_{i=1}^{n} \frac{\partial}{\partial x^i}(f \circ x^{-1})|_{x(p)} (\frac{d}{dt}(x \circ \rho)^i)_{t=t_0}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \frac{\partial}{\partial x^i}(f \circ x^{-1})|_{x(p)}(\frac{d}{dt}\hat{\rho}^i(t))_{t=t_0} \\
&= a\sum_{i=1}^{n} \frac{\partial}{\partial x^i}(f \circ x^{-1})|_{x(p)}(\frac{d}{dt}(x \circ \gamma)^i)_{t=t_0} \\
&+ b\sum_{i=1}^{n} \frac{\partial}{\partial x^i}(f \circ x^{-1})|_{x(p)}(\frac{d}{dt}(x \circ \gamma)^i)_{t=t_0} \\
&= a\dot{\gamma}_p(f) + b\dot{\sigma}_p(f)
\end{aligned}
\tag{2.11}
$$

So it follows that $a\dot{\gamma}_p + b\dot{\sigma}_p$ is the tangent vector to ρ at p.
In order to compute the dimension of the vector space it suffices to compute a basis. To do this, define n curves $\rho(i)$ for $i = 1, \ldots, n$ passing through p by

$$(x \circ \rho(i))(t)^j = (x(p))^j + t\delta_i^j \tag{2.12}$$

Using (2.9) it is straightforward to compute the tangent vectors to the curves $\rho(i)$ at p;

$$\dot{\rho}(i)_p(f) = \frac{\partial}{\partial x^i}(f \circ x^{-1})|_{x(p)} \tag{2.13}$$

and hence, if γ is a curve passing through p then (2.9) implies that

$$\dot{\gamma}_p(f) = \sum_{i=1}^{n} \dot{\rho}(i)_p(f)\dot{\gamma}_p^i \tag{2.14}$$

and hence it follows that $\dot{\gamma}_p = \sum_{i=1}^{n} \dot{\gamma}_p^i \dot{\rho}(i)_p$. Hence the tangent vectors to the curves $\rho(i)$ at p span $T_p(M)$. ■

Given the expression (2.13), it is conventional to write the tangent vectors to the curves $\rho(i)$ at p as

$$\dot{\rho}(i)_p = (\frac{\partial}{\partial x^i})_p \tag{2.15}$$

Lemma 1. *Suppose that M_1, M_2 are smooth manifolds of dimension n_1, n_2 respectively. Let $M = M_1 \times M_2$ be the Cartesian product manifold and suppose $p = (p_1, p_2) \in M$. Then $T_p(M) = T_{p_1}(M_1) \bigoplus T_{p_2}(M_2)$.*

Proof

Suppose $V_p \in T_p(M)$. Then V is the tangent vector to a smooth curve $\gamma(t)$, with $\gamma(t_0) = p$. Write $\gamma(t) = (\gamma_1(t), \gamma_2(t))$; $\gamma_i(t)$ is then a smooth curve in M_i and $\gamma_i(t_0) = p_i$ for $i = 1, 2$.

Let f be a smooth function $f : M \to \mathbb{R}$. Suppose that x^a are local co-ordinates on M_1 for $a = 1, \ldots, n_1$ and y^m are local co-ordinates on M_2 for $m = 1, \ldots, n_2$ corresponding to charts $U_1 \subset M_1$ and $U_2 \subset M_2$.

Then one has $n_1 + n_2$ local co-ordinates z^α on M where if $q = (q_1, q_2) \in U_1 \times U_2$,

$$z(q_1, q_2) = \big(x^1(q_1), \ldots, x^{n_1}(q_1), y^1(q_2), \ldots, y^{n_2}(q_2)\big) \tag{2.16}$$

Note that $f_1(q_1) = f(q_1, q_2)$ is a smooth function of q_1 when q_2 is fixed, and $f_2(q_2) = f(q_1, q_2)$ is a smooth function of q_2 when q_1 is fixed.

Then using the chain rule

$$\begin{aligned} V_p f &= \sum_{\alpha=1}^{n_1+n_2} \frac{\partial}{\partial z^\alpha}(f \circ z^{-1})|_{z(p)} \frac{d}{dt}((z \circ \gamma)^\alpha(t))|_{t=t_0} \\ &= \sum_{a=1}^{n_1} \frac{\partial}{\partial x^a}(f_1 \circ x^{-1})|_{(x(p_1),y(q_2))} \frac{d}{dt}((x \circ \gamma_1)^a(t))|_{t=t_0} \\ &+ \sum_{j=1}^{n_2} \frac{\partial}{\partial y^j}(f_2 \circ y^{-1})|_{(x(p_1),y(q_2))} \frac{d}{dt}((y \circ \gamma_2)^j(t))|_{t=t_0} \\ &= (V(1)_p + V(2)_p) f \end{aligned} \tag{2.17}$$

where $V(1)_p$ is the tangent vector to γ_1 at p, and $V(2)_p$ is the tangent vector to γ_2 at p. Hence $V_p = V(1)_p + V(2)_p$. Conversely, given two smooth curves $\gamma_1(t)$, $\gamma_2(t)$ in M_1, M_2 passing through p_1 and p_2 at $t = t_0$, with associated tangent vectors $V(1)_p$ and $V(2)_p$, one can construct the smooth curve $\gamma(t) = (\gamma_1(t), \gamma_2(t))$ in M passing through $p = (p_1, p_2)$ at $t = t_0$. Then (2.17) shows that $V(1)_p + V(2)_p$ can be written as $V_p \in T_p(M)$.

2.5 Vector Fields and Commutators

The tangent space of M, $T(M)$ consists of the union

$$T(M) = \bigcup_{p \in M} T_p(M) \tag{2.18}$$

A vector field V on M is a map $V : M \to T(M)$ such that $V(p) = V_p \in T_p(M)$.

Note that $T(M)$ is a vector space with addition and scalar multiplication defined by

$$(X + Y)(f) = X(f) + Y(f) \tag{2.19}$$

where $X, Y \in T(M)$ and $f : M \to \mathbb{R}$ is smooth, and

$$(\alpha X)(f) = \alpha X(f) \tag{2.20}$$

for constant $\alpha \in \mathbb{R}$.

At a point $p \in M$, one can decompose V_p into its components with respect to a particular chart as

$$V_p = V_p^i \left(\frac{\partial}{\partial x^i}\right)_p \tag{2.21}$$

It is conventional to write

$$V = V^i \left(\frac{\partial}{\partial x^i}\right) \tag{2.22}$$

where $V^i = (V \circ x^{-1})(x^i)$ are functions $\mathbb{R}^n \to \mathbb{R}$ and $(\frac{\partial}{\partial x^i})$ is a locally defined vector field which satisfies

$$(\frac{\partial}{\partial x^i})x^j = \delta_i^j \tag{2.23}$$

It follows that $T(M)$ is n-dimensional with a local basis given by the $(\frac{\partial}{\partial x^i})$. The vector field is called smooth if the functions V^i are smooth functions on $\mathbb{R}^n$.

Suppose now that f is a smooth function on M and that V, W are vector fields on M. Then note that Vf can be regarded as a function $M \to \mathbb{R}$ defined by

$$(Vf)(p) = V_p f \tag{2.24}$$

Hence one can act on Vf with W_p at some $p \in M$ to find

$$\begin{aligned} W_p(Vf) &= W_p^i(\frac{\partial}{\partial x^i})_p(Vf) \\ &= W_p^i(\frac{\partial}{\partial x^i})_p(V^j\frac{\partial}{\partial x^j}(f \circ x^{-1}))|_{x(p)} \\ &= W_p^i\frac{\partial V^j}{\partial x^i}|_{x(p)}(\frac{\partial}{\partial x^j}(f \circ x^{-1}))|_{x(p)} \\ &+ W_p^i V_p^j(\frac{\partial^2}{\partial x^i \partial x^j}(f \circ x^{-1}))|_{x(p)} \end{aligned} \tag{2.25}$$

The fact that there are second order derivatives acting on f means that we cannot write $W_p(Vf) = Z_p f$ for some vector field Z.

However, these second order derivatives can be removed by taking the difference

$$W_p(Vf) - V_p(Wf) = \big(W_p^i\frac{\partial V^j}{\partial x^i}|_{x(p)} - V_p^i\frac{\partial W^j}{\partial x^i}|_{x(p)}\big)(\frac{\partial}{\partial x^j}(f \circ x^{-1}))|_{x(p)} \tag{2.26}$$

which can be written as $Z_p f$ where Z is a vector field called the commutator or alternatively the Lie bracket of W and V which we denote by $[W, V]$ with components

$$[W, V]^j = W^i\frac{\partial V^j}{\partial x^i} - V^i\frac{\partial W^j}{\partial x^i} \tag{2.27}$$

<u>Exercise:</u>

Prove that the Lie bracket satisfies

i) Skew-symmetry: $[X, Y] = -[Y, X]$ for all smooth vector fields X, $Y \in T(M)$.

ii) Linearity: $[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z]$ for α, β constants and X, Y, $Z \in T(M)$.

iii) The Jacobi identity: $[[X, Y], Z] + [[Z, X], Y] + [[Y, Z], X] = 0$ for all X, Y, $Z \in T(M)$.

Definition 11. *Let V be a smooth vector field on M. An integral curve $\sigma(t)$ of V is a curve whose tangent vector at $\sigma(t)$ is $V|_{\sigma(t)}$, i.e.*

$$\frac{d}{dt}(\sigma^i(t)) = V^i_{\sigma(t)} \tag{2.28}$$

where in a slight abuse of notation, $\sigma^i(t) = (x \circ \sigma)^i(t)$ for some local co-ordinates x. Such a curve is guaranteed to exist and to be unique (at least locally, given an initial condition), by the standard existence and uniqueness theorems for ODE's.

2.6 Push-Forwards of Vector Fields

Suppose that M, N are two smooth manifolds and $f : M \to N$ is a smooth map. Then there is an induced map

$$f_* : T(M) \to T(N) \tag{2.29}$$

which maps the tangent vector of a curve γ passing through a point $p \in M$ to the tangent vector of the curve $f \circ \gamma$ passing through $f(p) \in N$.

In particular, for each smooth function h on N, and if γ is a curve passing through $p \in M$ with $\gamma(0) = p$, and if $V_p \in T_p(M)$ is the tangent vector of γ at p then $f_* V_p \in T_{f(p)}(N)$ is given by

$$\begin{aligned}(f_* V_p)h &= \frac{d}{dt}\big(h \circ (f \circ \gamma)\big)_{t=0} \\ &= V_p(h \circ f)\end{aligned} \tag{2.30}$$

Hence it is clear that the push-forward map f_* is linear on the space of tangent vectors.

Note that if M, N and Q are manifolds, and $f : M \to N$, $g : N \to Q$ are smooth functions then if $h : Q \to \mathbb{R}$ is smooth and $p \in M$,

$$\begin{aligned}\big((g \circ f)_* V_p\big)(h) &= V_p(h \circ (g \circ f)) \\ &= V_p((h \circ g) \circ f) \\ &= (f_* V_p)(h \circ g) \\ &= \big(g_*(f_* V_p)\big)(h)\end{aligned} \tag{2.31}$$

and hence

$$(g \circ f)_* = g_* \circ f_* \tag{2.32}$$

2.7 Left-Invariant Vector Fields

Suppose that G is a Lie group and a, $g \in G$. Define the operation of left-translation $L_a : G \to G$ by

$$L_a g = ag \tag{2.33}$$

L_a defined in this fashion is a differentiable invertible map from G onto G. Hence, one can construct the push-forward L_{a*} of vector fields on G with respect to L_a.

Definition 12. *A vector field $X \in T(G)$ is said to be left-invariant if*

$$L_{a*}(X|_g) = X|_{ag} \tag{2.34}$$

Given $v \in T_e(G)$ one can construct a unique left-invariant vector field $X(v) \in T(G)$ with the property that $X(v)_e = v$ using the push-forward by

$$X(v)|_g = L_{g*}v \tag{2.35}$$

To see that $X(v)$ is left-invariant, note that

$$X(v)|_{ag} = L_{(ag)*}v \tag{2.36}$$

but from (2.32) it follows that as $L_{ag} = L_a \circ L_g$ we must have

$$L_{(ag)*}v = (L_a \circ L_g)_* v = L_{a*}(L_{g*}v) = L_{a*}X(v)_g \tag{2.37}$$

so $X(v)$ is left-invariant. Hence there is a 1-1 correspondence between elements of the tangent space at e and the set of left-invariant vector fields.

Proposition 2. *The set of left-invariant vector fields is closed under the Lie bracket, i.e. if X, $Y \in T(G)$ are left-invariant then so is $[X,Y]$.*

Proof

Suppose that $f : G \to \mathbb{R}$ is a smooth function. Then

$$\begin{aligned}\big(L_{a*}[X,Y]_g\big)f &= [X,Y]_g(f \circ L_a) \\ &= X_g(Y(f \circ L_a)) - Y_g(X(f \circ L_a))\end{aligned} \tag{2.38}$$

But as X is left-invariant, $L_{a*}X_g = X_{ag}$ so

$$X_{ag}f = (L_{a*}X_g)f = X_g(f \circ L_a) \tag{2.39}$$

so replacing f with Yf in the above we find

$$X_g((Yf) \circ L_a) = X_{ag}(Yf) \tag{2.40}$$

Moreover, as Y is left-invariant, it is straightforward to show that

$$\begin{aligned}\big(Y(f \circ L_a)\big)g &= Y_g(f \circ L_a) \\ &= (L_{a*}Y_g)f \\ &= Y_{ag}(f) \\ &= (Yf)(ag) \\ &= \big((Yf) \circ L_a\big)g\end{aligned} \tag{2.41}$$

so $Y(f \circ L_a) = (Yf) \circ L_a$

Hence

$$\begin{aligned}X_g(Y(f \circ L_a)) - Y_g(X(f \circ L_a)) &= X_g((Yf) \circ L_a) - Y_g((Xf) \circ L_a) \\ &= X_{ag}(Yf) - Y_{ag}(Xf) \\ &= [X,Y]_{ag}f\end{aligned} \tag{2.42}$$

So $L_{a*}[X,Y]_g = [X,Y]_{ag}$, hence $[X,Y]$ is left-invariant. ■

2.8 Lie Algebras

Definition 13. *Suppose that G is a Lie group. Then the Lie algebra $\mathcal{L}(G)$ associated with G is $T_e(G)$, the tangent space of G at the origin, together with a Lie bracket $[\ ,\ ]: \mathcal{L}(G) \times \mathcal{L}(G) \to \mathcal{L}(G)$ which is defined by*

$$[v, w] = [L_* v, L_* w]_e \tag{2.43}$$

for $v, w \in T_e(G)$, $L_* v$ and $L_* w$ denote the smooth vector fields on G obtained by pushing forward v and w by left-multiplication (i.e. $L_* v|_g = L_{g*} v$), and $[L_* v, L_* w]$ is the standard vector field commutator. As the Lie bracket on $\mathcal{L}(G)$ is obtained from the commutator of vector fields, it follows that the Lie bracket is

i) Skew-symmetric: $[v, w] = -[w, v]$ for all $v, w \in \mathcal{L}(G)$.

ii) Linear: $[\alpha v_1 + \beta v_2, w] = \alpha[v_1, w] + \beta[v_2, w]$ for α, β constants and v_1, v_2, $w \in \mathcal{L}(G)$,

iii) and satisfies the Jacobi identity: $[[v, w], z] + [[z, v], w] + [[w, z], v] = 0$ for all v, w, $z \in \mathcal{L}(G)$.

where (ii) follows because the push forward map is linear on the space of vector fields, and (iii) follows because as a consequence of Proposition 2, $L_{g*}[v, w] = [L_* v, L_* w]_g$.

More generically, one can also define a Lie algebra to be a vector space $\mathfrak{g}$ equipped with a map $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying $(i), (ii), (iii)$ above.

Definition 14. *Suppose that $\{T_i : i = 1, \dots, n\}$ is a basis for $\mathcal{L}(G)$. Then the T_i are called generators of the Lie algebra. As $[T_i, T_j] \in \mathcal{L}(G)$ it follows that there are constants ${c_{ij}}^k$ such that*

$$[T_i, T_j] = {c_{ij}}^k T_k \tag{2.44}$$

The constants ${c_{ij}}^k$ are called the structure constants of the Lie algebra.

The structure constants are constrained by the antisymmetry of the Lie bracket to be antisymmetric in the first two indices;

$${c_{ij}}^k = -{c_{ji}}^k \tag{2.45}$$

Also, the Jacobi identity implies

$$[[T_i, T_j], T_k] + [[T_j, T_k], T_i] + [[T_k, T_i], T_j] = 0 \tag{2.46}$$

which gives an additional constraint on the structure constants

$${c_{ij}}^\ell {c_{\ell k}}^m + {c_{jk}}^\ell {c_{\ell i}}^m + {c_{ki}}^\ell {c_{\ell j}}^m = 0 \tag{2.47}$$

2.9 Matrix Lie Algebras

The Lie algebras of matrix Lie groups are of particular interest. Suppose that G is a matrix Lie group, and $V \in T(G)$ is a smooth vector field. Let f be a smooth function of the matrix components g^{AB}. Then if $h \in G$,

$$\begin{aligned} V_h f &= V_h^m \frac{\partial f}{\partial x^m} \\ &= V_h^m \frac{\partial g^{AB}}{\partial x^m} \frac{\partial f}{\partial g^{AB}} \\ &= V_h^{AB} \frac{\partial f}{\partial g^{AB}} \end{aligned} \tag{2.48}$$

where

$$V_h^{AB} = V_h^m \frac{\partial g^{AB}}{\partial x^m} \tag{2.49}$$

defines a tangent matrix associated with the components V_h^m of V at h. Each vector field has a corresponding tangent matrix, and it will often be most convenient to deal with these matrices instead of more abstract vector fields as differential operators.

In particular, if $\gamma(t)$ is some curve in G with tangent vector V then

$$\begin{aligned} Vf &= \frac{d}{dt}(f \circ \gamma(t)) \\ &= \frac{dg^{AB}}{dt} \frac{\partial f}{\partial g^{AB}} \end{aligned} \tag{2.50}$$

hence the tangent vector to the curve corresponds to the matrix $\frac{dg^{AB}}{dt}$. We will frequently denote the identity element of a matrix Lie group by $e = \mathbb{I}$

Examples of matrix Lie algebras are

- a) $GL(n,\mathbb{R})$: the co-ordinates of $GL(n,\mathbb{R})$ are the n^2 components of the matrices, so $GL(n,\mathbb{R})$ is n^2-dimensional. There is no restriction on tangent matrices to curves in $GL(n,\mathbb{R})$, the space of tangent vectors is $M_{n\times n}(\mathbb{R})$, the set of $n \times n$ real matrices.

- b) $GL(n,\mathbb{C})$: the co-ordinates of $GL(n,\mathbb{C})$ are the n^2 components of the matrices, so $GL(n,\mathbb{C})$ is $2n^2$-dimensional when viewed as a real manifold. There is no restriction on tangent matrices to curves in $GL(n,\mathbb{C})$, the space of tangent vectors is $M_{n\times n}(\mathbb{C})$, the set of $n \times n$ complex matrices.

- c) $SL(n,\mathbb{R})$: Suppose that $M(t)$ is a curve in $SL(n,\mathbb{R})$ with $M(0) = \mathbb{I}$. To compute the restrictions on the tangent vectors to the curve note that

$$\det M(t) = 1 \tag{2.51}$$

so, on differentiating with respect to t,

$$\mathrm{Tr}\big(M^{-1}(t)\frac{dM(t)}{dT}\big) = 0 \tag{2.52}$$

and so if we denote the tangent vector at the identity to be $m = \frac{dM(t)}{dt}|_{t=0}$ then Tr $m = 0$. The tangent vectors correspond to traceless matrices. Hence $SL(n, \mathbb{R})$ is $n^2 - 1$ dimensional.

- d) $O(n)$: suppose that $M(t)$ is a curve in $O(n)$ with $M(0) = \mathbb{I}$. To compute the restrictions on the tangent vectors to the curve note that

$$M(t)M(t)^T = 1 \tag{2.53}$$

so, on differentiating with respect to t,

$$\frac{dM(t)}{dt}M(t)^T + M(t)\frac{dM(t)^T}{dt} = 0 \tag{2.54}$$

and hence if $m = \frac{dM(t)}{dt}|_{t=0}$ then $m+m^T = 0$. The tangents to the curve at the identity correspond to antisymmetric matrices. There are $\frac{1}{2}n(n-1)$ linearly independent antisymmetric matrices, hence $O(n)$ is $\frac{1}{2}n(n-1)$-dimensional.

Note that the Lie algebra of $SO(2)$ is 1-dimensional and is spanned by

$$T_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.55}$$

As $[T_1, T_1] = 0$ it follows trivially that the Lie bracket vanishes

- f) $SO(n)$ the group of $n \times n$ real matrices such that if $M \in SO(n)$ then $MM^T = 1$ and $\det M = 1$. By the reasoning in (c) and (e) it follows that the tangent matrices at the identity are skew-symmetric matrices (these are automatically traceless).

  As the skew symmetric matrices are automatically traceless, it follows that the Lie algebra $\mathcal{L}(SO(n))$ of $SO(n)$ is identical to the Lie algebra of $O(n)$. If $v, w \in \mathcal{L}(SO(n))$ are skew-symmetric matrices it is straightforward to show that the matrix commutator $[v, w]$ is also skew symmetric, as $[v, w]^T = (vw - wv)^T = w^T v^T - v^T w^T = [w, v] = -[v, w]$. Hence $[v, w] \in \mathcal{L}(SO(n))$ as expected. We will show that vector field commutation can be reduced to tangent matrix commutation for matrix Lie groups.

Exercise

Show that the tangent vectors of $U(n)$ at $\mathbb{I}$ consist of antihermitian matrices, and the tangent vectors of $SU(n)$ at $\mathbb{I}$ are traceless antihermitian matrices.

Proposition 3. *Suppose that G is a matrix Lie group and V is a smooth vector field on G and $a \in G$ is fixed. If $\hat{V}$ denotes the tangent matrix associated with V, then the tangent matrix associated with the push-forward $L_{a*}V$ is $a\hat{V}$.*

Proof

Suppose $h \in G$, and $f : G \to \mathbb{R}$ is a smooth function on G. Consider the tangent vector $L_{a*}V_h$ defined at ah

Then

$$\begin{aligned}(L_{a*}V_h)f &= V_h(f \circ L_a) \\ &= V_h\tilde{f}\end{aligned} \tag{2.56}$$

where $\tilde{f}(g) = f(ag)$.

So

$$\begin{aligned}(L_{a*}V_h)f &= V_h^{AB}\frac{\partial \tilde{f}}{\partial g^{AB}}|_h \\ &= V_h^{AB}\frac{\partial f}{\partial g^{PQ}}|_{ah}\frac{\partial}{\partial g^{AB}}((ag)^{PQ}) \\ &= V_h^{AB}\frac{\partial f}{\partial g^{PB}}|_{ah}a^{PA} \\ &= (a\hat{V})^{AB}\frac{\partial f}{\partial g^{AB}}|_{ah}\end{aligned} \tag{2.57}$$

So it follows that the tangent matrix associated with $L_{a*}V_h$ is $a\hat{V}$. ■

Using this result, it is possible to re-interpret the commutator of two left-invariant vector fields in terms of the matrix commutators of their associated matrices.

Proposition 4. *Suppose that G is a matrix Lie group and that $v, w \in T_e(G)$ and V, W are the left-invariant vector fields defined by $V_g = L_{g*}v$, $W_g = L_{g*}w$. Then the matrix associated with $[V, W]_e$ is the matrix commutator of $[\hat{v}, \hat{w}]$ where $\hat{v}$ and $\hat{w}$ are the matrices associated with v and w.*

Proof

By definition, the matrix associated with $[V, W]$ is

$$\begin{aligned}[V,W]^{AB} &= [V,W]^m\frac{\partial g^{AB}}{\partial x^m} \\ &= V^p\frac{\partial W^m}{\partial x^p}\frac{\partial g^{AB}}{\partial x^m} - W^p\frac{\partial V^m}{\partial x^p}\frac{\partial g^{AB}}{\partial x^m} \\ &= V^p\frac{\partial}{\partial x^p}\Big(W^m\frac{\partial g^{AB}}{\partial x^m}\Big) - W^p\frac{\partial}{\partial x^p}\Big(V^m\frac{\partial g^{AB}}{\partial x^m}\Big) \\ &= V^p\frac{\partial \hat{W}^{AB}}{\partial x^p} - W^p\frac{\partial \hat{V}^{AB}}{\partial x^p}\end{aligned} \tag{2.58}$$

where $\hat{V}$ and $\hat{W}$ denote the matrices associated with V and W. But from the previous proposition $\hat{V_g}^{AB} = g^{AC}\hat{v}^{CB}$ and $\hat{W_g}^{AB} = g^{AC}\hat{w}^{CB}$ so

$$
\begin{aligned}
[V,W]_e^{AB} &= V^P \frac{\partial g^{AC}}{\partial x^p}|_e \hat{w}^{CB} - W^P \frac{\partial g^{AC}}{\partial x^p}|_e \hat{v}^{CB} \\
&= \hat{v}^{AC}\hat{w}^{CB} - \hat{w}^{AC}\hat{v}^{CB} \\
&= [\hat{v},\hat{w}]^{AB}
\end{aligned}
\tag{2.59}
$$

as required. ■

We have therefore shown that if G is a matrix Lie group then the elements $\mathcal{L}(G)$ can be associated with matrices and the Lie bracket is then simply standard matrix commutation by Proposition 4 (which can be directly checked satisfies all three of the Lie bracket for Lie algebras). In the literature, it is often conventional to denote the Lie algebra of $SO(n)$ by $\mathfrak{so}(n)$, $\mathfrak{su}(n)$ is the Lie algebra of $SU(n)$, $\mathfrak{u}(n)$ the Lie algebra for $U(n)$ etc. We will however continue to use the notation $\mathcal{L}(G)$ for the Lie algebra of Lie group G.

Observe that the image $[\mathcal{L}(G),\mathcal{L}(G)]$ under the Lie bracket need not be the whole of $\mathcal{L}(G)$. This is clear for $SO(2)$, as the Lie bracket vanishes identically in that case. Recall that the Lie bracket on $\mathbb{R}$ viewed as a Lie group under addition vanishes identically as well. If G is a connected 1-dimensional Lie group then G must either be isomorphic to $\mathbb{R}$ or $SO(2)$ (equivalently $U(1)$).

2.10 One Parameter Subgroups

Definition 15. *A curve $\sigma:\mathbb{R}\to G$ is called a one-parameter subgroup if $\sigma(s)\sigma(t)=\sigma(s+t)$ for all $s,t\in\mathbb{R}$.*

Note that if $\sigma(t)$ is a 1-parameter subgroup then $\sigma(0)=e$.

We shall show that these subgroups arise naturally as integral curves of left-invariant vector fields.

Proposition 5. *Suppose that V is a left-invariant vector field. Let $\sigma(t)$ be the integral curve of V which passes through e when $t=0$.*

Then $\sigma(t)$ is a 1-parameter subgroup of G.

<u>Proof</u>

Let x denote some local co-ordinates.

Consider the curves $\chi_1(t)=\sigma(s)\sigma(t)$ and $\chi_2(t)=\sigma(s+t)$ for fixed s.

These satisfy the same initial conditions $\chi_1(0)=\chi_2(0)=\sigma(s)$.

By definition, χ_2 satisfies the ODE

$$
\begin{aligned}
\frac{d}{dt}((x\circ\chi_2(t))^n) &= \frac{d}{d(s+t)}((x\circ\sigma(s+t))^n) \\
&= V_{\sigma(s+t)}(x^n) \\
&= V_{\chi_2(t)}(x^n)
\end{aligned}
\tag{2.60}
$$

Consider

$$
\frac{d}{dt}((x\circ\chi_1(t))^n) = \frac{d}{dt}\big((x\circ L_{\sigma(s)}\circ\sigma(t))^n\big)
$$

$$
\begin{aligned}
&= \frac{d}{dt}\big(((x\circ L_{\sigma(s)}\circ x^{-1})\circ(x\circ\sigma)(t))^n\big)\\
&= \frac{\partial}{\partial x^m}\big((x\circ L_{\sigma(s)}\circ x^{-1})^n\big)|_{x\circ\sigma(t)}\frac{d}{dt}((x\circ\sigma(t))^m)
\end{aligned}
\tag{2.61}
$$

where we have used the chain rule. But by definition of $\sigma(t)$,

$$
\frac{d}{dt}((x\circ\sigma(t))^m) = V^m_{\sigma(t)}
\tag{2.62}
$$

Hence, substituting this into the above:

$$
\begin{aligned}
\frac{d}{dt}((x\circ\chi_1(t))^n) &= V^m_{\sigma(t)}\frac{\partial}{\partial x^m}\big((x\circ L_{\sigma(s)}\circ x^{-1})^n\big)|_{x\circ\sigma(t)}\\
&= V_{\sigma(t)}\big((x\circ L_\sigma(s))^n\big)\\
&= (L_{\sigma(s)*}V_{\sigma(t)})(x^n) \qquad \text{(by definition of push}-\text{forward)}\\
&= V_{\chi_1(t)}(x^n) \qquad \text{(as V is left}-\text{invariant)}
\end{aligned}
\tag{2.63}
$$

So χ_1, χ_2 satisfy the same ODE with the same initial conditions.

Hence it follows that $\sigma(s)\sigma(t)=\sigma(s+t)$, i.e. σ defines a 1-parameter subgroup. ■

The converse is also true: a 1-parameter subgroup $\sigma(t)$ has left-invariant tangent vectors

Proposition 6. *Suppose $\sigma(t)$ is a 1-parameter subgroup of G with tangent vector V. Suppose $V_e = v$. Then $V_{\sigma(t)} = L_{\sigma(t)*}v$, i.e. the tangent vectors are obtained by pushing forward the tangent vector at the identity.*

Proof

Suppose $f: G \to \mathbb{R}$ is a smooth function. Then

$$
\begin{aligned}
V_{\sigma(t)}f &= \frac{d}{dt}\big((f\circ\sigma)(t)\big)\\
&= \lim_{h\to 0}\Big(\frac{f(\sigma(t+h))-f(\sigma(t))}{h}\Big)\\
&= \lim_{h\to 0}\Big(\frac{f(\sigma(t)\sigma(h))-f(\sigma(t))}{h}\Big)\\
&= \frac{d}{dt'}(f\circ L_{\sigma(t)}\circ\sigma(t'))|_{t'=0}\\
&= (L_{\sigma(t)*}v)f
\end{aligned}
\tag{2.64}
$$

so $V_{\sigma(t)} = L_{\sigma(t)*}v$. ■

From this we obtain the corollory

Corollory 1. *Suppose that $\sigma(t)$, $\mu(t)$ are two 1-parameter subgroups of G with tangent vectors V, W respectively, with $V_e = W_e = u$. Then $\sigma(t) = \mu(t)$ for all t.*

Proof

Note that

$$
\frac{d}{dt}\big((x\circ\sigma(t))^n\big) = V_{\sigma(t)}x^n
$$

$$= (L_{\sigma(t)*}u)x^n \tag{2.65}$$

and also

$$\begin{aligned}\frac{d}{dt}\big((x \circ \mu(t))^n\big) &= W_{\sigma(t)}x^n \\ &= (L_{\mu(t)*}u)x^n \end{aligned} \tag{2.66}$$

So $x \circ \sigma$ and $x \circ \mu$ satisfy the same ODE and with the same initial conditions, hence $\sigma(t) = \mu(t)$. ■

2.11 Exponentiation

Definition 16. *Suppose $v \in T_e(G)$, Then we define the exponential map* $\exp : T_e(G) \to G$ *by*

$$\exp(v) = \sigma_v(1) \tag{2.67}$$

where $\sigma_v(t)$ denotes the 1-parameter subgroup generated by $X(v)$, and $X(v)$ is the left-invariant vector field obtained via the push-forward $X(v)_g = L_{g*}v$

Note that $\exp(0) = e$.

Proposition 7. *If $v \in T_e(G)$ and $t \in \mathbb{R}$ then*

$$\exp(tv) = \sigma_v(t) \tag{2.68}$$

and hence $\exp((t_1 + t_2)v) = \exp(t_1 v)\exp(t_2 v)$.

Proof

Take $a \in \mathbb{R}$, $a \neq 0$. Note that $\sigma_v(at)$ and $\sigma_{av}(t)$ are both 1-parameter subgroups of G. The tangent vector to $\sigma_{av}(t)$ at the origin is av.

We also compute the tangent vector to $\sigma_v(at)$ at e via

$$\frac{d}{dt}\big((x \circ \sigma(at))^n\big)_{t=0} = a\frac{d}{d(at)}\big((x \circ \sigma(at))^n\big)_{at=0} = av^n \tag{2.69}$$

So $\sigma_v(at)$ and $\sigma_{av}(t)$ have the same tangent vector av at the origin. Therefore $\sigma_v(at) = \sigma_{av}(t)$.

Hence

$$\exp(tv) = \sigma_{tv}(1) = \sigma_v(t) \tag{2.70}$$

as required. ■

2.12 Exponentiation on matrix Lie groups

Suppose that G is a matrix Lie group, and $v \in T_e(G)$ is some tangent matrix. The exponential $\exp(tv)$ produces a curve in G with $\frac{d}{dt}(\exp(tv))|_{t=0} = v$ satisfying $\exp((t_1 + t_2)v) = \exp(t_1 v)\exp(t_2 v)$

It is then straightforward to show that

$$\begin{aligned} \frac{d}{dt}(\exp(tv))|_{t=t_0} &= \lim_{t\to 0}\left(t^{-1}(\exp((t_0+t)v) - \exp(t_0 v))\right) \\ &= \lim_{t\to 0}\left(t^{-1}(\exp(tv) - \mathbb{I})\exp(t_0 v)\right) \\ &= v\exp(t_0 v) \end{aligned} \tag{2.71}$$

Similarly, one also finds $\frac{d}{dt}(\exp(tv))|_{t=t_0} = \exp(t_0 v)v$, so v commutes with $\exp(tv)$.

It is clear that $\frac{d}{dt}\exp(tv) = v\exp(tv)$ implies that $\exp(tv)$ is infinitely differentiable (as expected as the integral curve is smooth by construction). Then by elementary analysis, one can compute the power series expansion for $\exp(tv)$ as

$$\exp(tv) = \sum_{n=0}^{\infty} \frac{t^n v^n}{n!} \tag{2.72}$$

with a remainder term which converges to 0 (with respect to the supremum norm on matrices, for example). Hence, for matrix Lie groups, the Lie group exponential operator corresponds to the usual operation of matrix exponentiation.

Comment: Suppose that G_1 and G_2 are Lie groups. Then $G = G_1 \times G_2$ is a Lie group, and by Lemma 1, $\mathcal{L}(G) = \mathcal{L}(G_1) \bigoplus \mathcal{L}(G_2)$.

Conversely, suppose Lie groups G, G_1, G_2 are such that $\mathcal{L}(G) = \mathcal{L}(G_1) \bigoplus \mathcal{L}(G_2)$. Then by exponentiation, it follows that, at least in a local neighbourhood of e, G has the local geometric structure of $G_1 \times G_2$. However, as it is not in general possible to reconstruct the whole group in this fashion, one cannot say that $G = G_1 \times G_2$ globally (typically there will be some periodic identification somewhere in the Cartesian product group).

In general, one cannot reconstruct the entire Lie group by exponentiating elements of the Lie algebra. Consider for example, $SO(2)$ and $O(2)$. Both $\mathcal{L}(O(2))$ and $\mathcal{L}(SO(2))$ are generated by

$$T_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.73}$$

however it is straightforward to show that

$$e^{\theta T_1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{2.74}$$

which always has determinant +1. So $SO(2) = \exp(\mathcal{L}(SO(2)))$ but $O(2) \neq \exp(\mathcal{L}(O(2)))$. However, there do exist neighbourhoods B_0 of $0 \in \mathcal{L}(G)$ and B_1 of $\mathbb{I} \in G$ such that the map $\exp : B_0 \to B_1$ is invertible. (The inverse is called log by convention).

2.13 Integration on Lie Groups

Suppose that G is a matrix Lie group, and let V be a left-invariant vector field on G, and suppose that the associated tangent matrix to V at the identity is $\hat{v}$.

Then if x are some local co-ordinates on G, we know that

$$g(x)\hat{v} = V^m_{g(x)} \frac{\partial g(x)}{\partial x^m} \tag{2.75}$$

From this formula, it is clear that if $h \in G$ is a constant matrix then

$$V^m_{g(x)} = V^m_{hg(x)} \tag{2.76}$$

If $H = \{h_1, \ldots, h_r\}$ is a finite group, and $f : H \to \mathbb{R}$ is a function, then the integral of f over H is simply

$$\sum_{i=1}^{r} f(h_i) \tag{2.77}$$

and note that if $h \in H$ is fixed then

$$\sum_{i=1}^{r} f(h_i) = \sum_{i=1}^{r} f(hh_i) \tag{2.78}$$

We wish to construct an analogous integral over a matrix Lie group G. Suppose that x, y are co-ordinates on G and define

$$d^n x = dx^1 \ldots dx^n, \qquad d^n y = dy^1 \ldots dy^n \tag{2.79}$$

Note that $d^n x$ and $d^n y$ are related by

$$d^n x = J^{-1} d^n y \tag{2.80}$$

where J is the Jacobian $J = \det\left(\frac{\partial y^i}{\partial x^j}\right)$.

Now suppose that μ_i for $i = 1, \ldots, n$ is a basis of left-invariant vector fields. Then

$$\mu_i|_{g(x)} = \mu^j_{i,g(x)} \frac{\partial}{\partial x^j} \tag{2.81}$$

Then we have

$$\mu^j_{i,g(x)} \frac{\partial}{\partial x^j} = \mu^j_{i,g(x)} \frac{\partial y^k}{\partial x^j} \frac{\partial}{\partial y^k} = \mu^j_{i,g(y)} \frac{\partial}{\partial y^j} \tag{2.82}$$

so

$$\mu^j_{i,g(y)} = \mu^k_{i,g(x)} \frac{\partial y^j}{\partial x^k} \tag{2.83}$$

and hence

$$\det\left(\mu^j_{i,g(x)}\right) = J^{-1} \det\left(\mu^j_{i,g(y)}\right) \tag{2.84}$$

Motivated by this, we make the

Definition 17. *The Haar measure is defined by*

$$d^n x \big(\det(\mu^j_{i,g(x)}) \big)^{-1} \tag{2.85}$$

Then by the previous reasoning,

$$d^n x \big(\det(\mu^j_{i,g(x)}) \big)^{-1} = d^n y \big(\det(\mu^j_{i,g(y)}) \big)^{-1} \tag{2.86}$$

so the measure is invariant under changes of co-ordinates.

Also, if h is a constant matrix, then as the μ_i are left-invariant, $\mu^j_{i,g(x)} = \mu^j_{i,hg(x)}$, and so

$$d^n x \big(\det(\mu^j_{i,g(x)}) \big)^{-1} = d^n x \big(\det(\mu^j_{i,hg(x)}) \big)^{-1} \tag{2.87}$$

It follows that if $f : G \to \mathbb{R}$, then

$$\int d^n x \big(\det(\mu^j_{i,g(x)}) \big)^{-1} f(g(x)) = \int d^n x \big(\det(\mu^j_{i,g(x)}) \big)^{-1} f(hg(x)) \tag{2.88}$$

It can be shown that the Haar measure (up to multiplication by a non-zero constant) is the unique measure with this property.

Example: $SL(2,\mathbb{R})$

Consider $g \in SL(2,\mathbb{R})$,

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{2.89}$$

for $a, b, c, d \in \mathbb{R}$ constrained by $ad - bc = 1$. Note that

$$g^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{2.90}$$

We take co-ordinates $x^1 = b$, $x^2 = c$, $x^3 = d$ (in some neighbourhood of the identity). Then

$$g^{-1}\frac{\partial g}{\partial x^1} = \begin{pmatrix} c & d \\ -\frac{c^2}{d} & -c \end{pmatrix}, \qquad g^{-1}\frac{\partial g}{\partial x^2} = \begin{pmatrix} 0 & 0 \\ \frac{1}{d} & 0 \end{pmatrix}, \qquad g^{-1}\frac{\partial g}{\partial x^3} = \begin{pmatrix} -a & -b \\ \frac{ac}{d} & a \end{pmatrix} \tag{2.91}$$

Take

$$v_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{2.92}$$

to be a basis for $\mathcal{L}(SL(2,\mathbb{R}))$. Then note that

$$v_1 = -bg^{-1}\frac{\partial g}{\partial x^1} + cg^{-1}\frac{\partial g}{\partial x^2} - dg^{-1}\frac{\partial g}{\partial x^3}$$

$$
\begin{aligned}
v_2 &= dg^{-1}\frac{\partial g}{\partial x^2} \\
v_3 &= ag^{-1}\frac{\partial g}{\partial x^1} + cg^{-1}\frac{\partial g}{\partial x^3}
\end{aligned} \tag{2.93}
$$

It follows that the left-invariant vector fields obtained from pushing-forward the vector fields associated with v_1, v_2, v_3 at the identity with L_* are

$$
\begin{aligned}
\mu_1 &= -b\frac{\partial}{\partial x^1} + c\frac{\partial}{\partial x^2} - d\frac{\partial}{\partial x^3} \\
\mu_2 &= d\frac{\partial}{\partial x^2} \\
\mu_3 &= a\frac{\partial}{\partial x^1} + c\frac{\partial}{\partial x^3}
\end{aligned} \tag{2.94}
$$

So the matrix μ_i^j is

$$
(\mu_i^j) = \begin{pmatrix} -b & c & -d \\ 0 & d & 0 \\ a & 0 & c \end{pmatrix} \tag{2.95}
$$

As $\det(\mu_i^j) = d$ it follows that the Haar measure in these co-ordinates is $\frac{1}{d}\, db\, dc\, dd$.

2.14 Representations of Lie Groups

Definition 18. *Let V be a finite dimensional vector space (over $\mathbb{R}$ or $\mathbb{C}$) and let $GL(V)$ denote the space of invertible linear transformations $V \to V$. Then a representation of a Lie group G acting on V is a map $\mathcal{D} : G \to GL(V)$ such that*

$$
\mathcal{D}(g_1 g_2) = \mathcal{D}(g_1)\mathcal{D}(g_2) \tag{2.96}
$$

for all $g_2, g_2 \in G$. (i.e. $\mathcal{D}$ is a homomorphism). The dimension of the representation is given by $\dim \mathcal{D} = \dim V$.

Lemma 2. *If $\mathcal{D}$ is a representation of G then $\mathcal{D}(e) = 1$ where $1 \in GL(V)$ is the identity transformation, and if $g \in G$ then $\mathcal{D}(g^{-1}) = (\mathcal{D}(g))^{-1}$.*

Proof

Note that $\mathcal{D}(e) = \mathcal{D}(ee) = \mathcal{D}(e)\mathcal{D}(e)$ and so it follows that $\mathcal{D}(e) = 1$ where $1 \in GL(V)$ is the identity transformation.

If $g \in G$ then $1 = \mathcal{D}(e) = \mathcal{D}(gg^{-1}) = \mathcal{D}(g)\mathcal{D}(g^{-1})$, so $\mathcal{D}(g^{-1}) = (\mathcal{D}(g))^{-1}$. ■

If $M(V)$ denotes the set of all linear transformations $V \to V$, and $\mathcal{D} : G \to M(V)$ satisfies $\mathcal{D}(e) = 1$ together with the condition (2.96) then it follows from the reasoning used in the Lemma above that $\mathcal{D}(g)$ is invertible for all $g \in G$, with inverse $\mathcal{D}(g^{-1})$, and hence $\mathcal{D}$ is a representation.

We next define some useful representations

Definition 19. *The trivial representation is defined by $\mathcal{D}(g) = 1$ for all $g \in G$*

Definition 20. *If G is a matrix Lie group which is a subgroup of $GL(n,\mathbb{R})$ or $GL(n,\mathbb{C})$ then the group elements themselves act directly on n-component vectors. The fundamental representation is then defined by $\mathcal{D}(g) = g$.*

Definition 21. *If G is a matrix Lie group then the adjoint representation $Ad : G \to GL(\mathcal{L}(G))$ is defined by*

$$(Ad(g))X = gXg^{-1} \tag{2.97}$$

for $g \in G$ and $X \in \mathcal{L}(G)$ is a tangent matrix.

Lemma 3. *Ad(g) as defined above is a representation*
Proof

We first verify that if $X \in \mathcal{L}(g)$ then $Ad(g)X \in \mathcal{L}(G)$.

Fix $g \in G$. Next, recall that if $X \in \mathcal{L}(G)$ then there is some smooth curve in G, $\gamma(t)$ such that $X = \frac{d\gamma(t)}{dt}|_{t=0}$. Define a new smooth curve in G by $\rho(t) = g\gamma(t)g^{-1}$, then the tangent matrix to $\rho(t)$ at $t = 0$ is given by $\frac{d\rho(t)}{dt}|_{t=0} = g\frac{d\gamma(t)}{dt}|_{t=0}g^{-1} = gXg^{-1}$.

Hence $Ad(g)X \in \mathcal{L}(G)$.

It is clear that $Ad(g)$ is a linear transformation on $X \in \mathcal{L}(G)$.

Note that $Ad(e)X = eXe^{-1} = X$, so $Ad(e) = 1$. Also, if $g_1, g_2 \in G$ then

$$\begin{aligned} Ad(g_1g_2)X &= (g_1g_2)X(g_1g_2)^{-1} \\ &= g_1g_2Xg_2^{-1}g_1^{-1} \\ &= g_1(g_2Xg_2^{-1})g_1^{-1} \\ &= g_1(Ad(g_2)X)g_1^{-1} \\ &= Ad(g_1)Ad(g_2)X \end{aligned} \tag{2.98}$$

hence $Ad(g_1g_2) = Ad(g_1)Ad(g_2)$.

It then follows that $Ad(g)Ad(g^{-1}) = Ad(gg^{-1}) = Ad(e) = 1$ so $Ad(g)$ is invertible. ■

Definition 22. *Suppose $\mathcal{D}$ is a representation of G acting on V. A subspace $W \subset V$ is called an invariant subspace if $\mathcal{D}(g)w \in W$ for all $g \in G$ and $w \in W$.*

Definition 23. *Suppose $\mathcal{D}$ is a representation of G acting on V. Then $\mathcal{D}$ is reducible if there is an invariant subspace W of V with $W \neq 0$ and $W \neq V$. If the only invariant subspaces of V are 0 and V then $\mathcal{D}$ is called irreducible.*

Definition 24. *A representation $\mathcal{D}$ is called totally reducible if there exists a direct sum decomposition of V into subspaces W_i, $V = W_1 \bigoplus W_2 \bigoplus ... \bigoplus W_k$ where the W_i are invariant subspaces with respect to $\mathcal{D}$ and $\mathcal{D}$ restricted to W_i is irreducible.*

In terms of matrices, if $\mathcal{D}$ is totally reducible, then there is some basis of V in which $\mathcal{D}$ has a block diagonal form

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & 0 & \\ 0 & \mathcal{D}_2(g) & \\ & & \ddots \end{pmatrix} \tag{2.99}$$

$\mathcal{D}_i(g)$ denotes $\mathcal{D}(g)$ restricted to W_i.

Definition 25. *A representation $\mathcal{D}$ of G acting on V is faithful if $\mathcal{D}(g) = 1$ only if $g = e$.*

Definition 26. *Suppose that $\mathcal{D}$ is a representation of G acting on V where V is a vector space over $\mathbb{C}$ equipped with an inner product. Then $\mathcal{D}$ is a unitary representation if $\mathcal{D}(g) : V \to V$ satisfies $\mathcal{D}(g)\mathcal{D}(g)^\dagger = 1$ for all $g \in G$.*

Proposition 8. *A finite dimensional unitary representation is totally reducible*

Proof

If $\mathcal{D}$ is irreducible then we are done. Otherwise, suppose that W is an invariant subspace. Write $V = W \oplus W_\perp$. Suppose $v \in W_\perp$. Then if $w \in W$ and $g \in g$,

$$\begin{aligned} \langle \mathcal{D}(g)v, w\rangle &= \langle v, \mathcal{D}(g)^\dagger w\rangle \\ &= \langle v, \mathcal{D}(g)^{-1} w\rangle \\ &= \langle v, \mathcal{D}(g^{-1}) w\rangle \\ &= 0 \end{aligned} \tag{2.100}$$

as $D(g^{-1})w \in W$ because W is an invariant subspace.

Hence it follows that if $v \in W_\perp$ then $D(g)v \in W_\perp$ and so $W_\perp$ is also an invariant subspace. Repeating this process by considering $\mathcal{D}$ restricted to the invariant subspaces W and $W_\perp$ one obtains a direct sum decomposition of V into invariant (orthogonal) subspaces W_i such that $\mathcal{D}$ restricted to W_i is irreducible. ■

Proposition 9. *Let V_1, V_2 be finite dimensional vector spaces. Suppose $\mathcal{D}$ is a representation of G acting on V_1, and $A : V_1 \to V_2$ is an invertible linear transformation. Define $\tilde{\mathcal{D}}(g) = A\mathcal{D}(g)A^{-1}$. Then $\tilde{\mathcal{D}}$ is a representation of G on V_2.*

Proof

As $\tilde{\mathcal{D}}$ is a composition of invertible linear transformations, $\tilde{\mathcal{D}}$ is also an invertible linear transformation on V_2.

Also, if $g_1, g_2 \in G$

$$\begin{aligned} \tilde{\mathcal{D}}(g_1 g_2) &= A\mathcal{D}(g_1 g_2)A^{-1} \\ &= A\mathcal{D}(g_1)\mathcal{D}(g_2)A^{-1} \\ &= A\mathcal{D}(g_1)A^{-1}A\mathcal{D}(g_2)A^{-1} \\ &= \tilde{\mathcal{D}}(g_1)\tilde{\mathcal{D}}(g_2) \end{aligned} \tag{2.101}$$

and hence $\tilde{\mathcal{D}}$ is also a representation. ■

Definition 27. *Suppose $\mathcal{D}$ is a representation of G acting on V_1, and $A : V_1 \to V_2$ is an invertible linear transformation. Define $\tilde{\mathcal{D}}(g) = A\mathcal{D}(g)A^{-1}$. Then $\mathcal{D}$ and $\tilde{\mathcal{D}}$ are said to be equivalent representations.*

Proposition 10. *Schur's First Lemma: Suppose that $\mathcal{D}_1$ and $\mathcal{D}_2$ are two irreducible representations of G acting on V_1 and V_2 respectively and there exists a linear transformation $A : V_1 \to V_2$ such that*

$$A\mathcal{D}_1(g) = \mathcal{D}_2(g)A \tag{2.102}$$

for all $g \in G$. Then either $\mathcal{D}_1$ and $\mathcal{D}_2$ are equivalent representations, or $A = 0$.

Proof First note that

$$\text{Ker } A = \{\psi \in V_1 : A\psi = 0\} \tag{2.103}$$

is an invariant subspace of $\mathcal{D}_1$, because if $\psi \in \text{Ker } A$ then

$$A\mathcal{D}_1(g)\psi = \mathcal{D}_2(g)A\psi = 0 \tag{2.104}$$

so $\mathcal{D}_1(g)\psi \in \text{Ker } A$ for all $g \in G$. But $\mathcal{D}_1$ is irreducible on V_1, so one must have $\text{Ker } A = 0$ or $\text{Ker } A = V_1$, so A is 1-1 or $A = 0$.

Similarly,

$$\text{Im } A = \{\phi \in V_2 : \phi = A\psi \text{ for some } \psi \in V_1\} \tag{2.105}$$

is an invariant subspace of $\mathcal{D}_2$, because if $\phi \in \text{Im } A$ then there is some $\psi \in V_1$ such that $\phi = A\psi$ and hence

$$\mathcal{D}_2(g)\phi = \mathcal{D}_2(g)A\psi = A\mathcal{D}_1(g)\psi \tag{2.106}$$

and hence $\mathcal{D}_2(g)\phi \in \text{Im A}$ for all $g \in G$. But $\mathcal{D}_2$ is irreducible on V_2, so one must have $\text{Im } A = 0$ or $\text{Im } A = V_2$, i.e. $A = 0$ or A is onto.

Hence either $A = 0$ or A is both 1-1 and onto i.e. A is invertible. If A is invertible then $\mathcal{D}_1$ and $\mathcal{D}_2$ are equivalent. ■

Proposition 11. *Schur's Second Lemma: Suppose that $\mathcal{D}$ is an irreducible representation of G on V, where V is a vector space over $\mathbb{C}$, and $A : V \to V$ is a linear transformation such that*

$$A\mathcal{D}(g) = \mathcal{D}(g)A \tag{2.107}$$

for all $g \in G$. Then $A = \lambda 1$ for some $\lambda \in \mathbb{C}$.

Proof

As V is over $\mathbb{C}$, A has at least one eigenvalue. Let $\lambda \in \mathbb{C}$ be an eigenvalue of A, with corresponding eigenspace U ($U \neq 0$). Then U is an invariant subspace of V with respect to $\mathcal{D}$, for if $\psi \in U$ then

$$A\psi = \lambda\psi \tag{2.108}$$

and if $g \in G$, then

$$A\mathcal{D}(g)\psi = \mathcal{D}(g)A\psi = \mathcal{D}(g)(\lambda\psi) = \lambda\mathcal{D}(g)\psi \tag{2.109}$$

so $\mathcal{D}(g)\psi \in U$.

But $\mathcal{D}$ is irreducible on V, so this implies $U = V$ (as $U \neq 0$).

Hence it follows that $A = \lambda 1$. ■

Definition 28. *Suppose that $\mathcal{D}_1$ and $\mathcal{D}_2$ are representations of the Lie group G over vector spaces V_1 and V_2 respectively. Let $V = V_1 \bigotimes V_2$ be the standard tensor product vector space of V_1 and V_2 consisting of elements $v_1 \otimes v_2$ ($v_1 \in V_1$ and $v_2 \in V_2$) in the vector space dual to the space of bilinear forms on $V_1 \times V_2$. If $v_1 \otimes v_2 \in V$ then $v_1 \otimes v_2$ acts linearly on bilinear forms Ω via $v_1 \otimes v_2 \Omega = \Omega(v_1, v_2)$. V is equipped with pointwise addition and scalar multiplication which satisfy $(v_1 + w_1) \otimes (v_2 + w_2) = v_1 \otimes v_2 + v_1 \otimes w_2 + w_1 \otimes v_2 + w_1 \otimes w_2$ and $\alpha(v_1 \otimes v_2) = (\alpha v_1) \otimes v_2 = v_1 \otimes (\alpha v_2)$.*

Then the tensor product representation $\mathcal{D}$ is defined as a linear map on V satisfying

$$\mathcal{D}(g) v_1 \otimes v_2 = \mathcal{D}_1(g) v_1 \otimes \mathcal{D}_2(g) v_2 \tag{2.110}$$

for $g \in G$ and $v_1 \in V_1$ and $v_2 \in V_2$

Proposition 12. *The tensor product representation defined above is a representation.*
Proof

The map $\mathcal{D}(g)$ is linear by construction, also

$$\mathcal{D}(e) v_1 \otimes v_2 = \mathcal{D}_1(e) v_1 \otimes \mathcal{D}_2(e) v_2 = v_1 \otimes v_2 \tag{2.111}$$

because $\mathcal{D}_1(e) = 1$ and $\mathcal{D}_2(e) = 1$. Hence $\mathcal{D}(e) = 1$. And if $g_1, g_2 \in G$ then

$$\begin{aligned}
\mathcal{D}(g_1 g_2) v_1 \otimes v_2 &= \mathcal{D}_1(g_1 g_2) v_1 \otimes \mathcal{D}_2(g_1 g_2) v_2 \\
&= \mathcal{D}_1(g_1)\mathcal{D}_1(g_2) v_1 \otimes \mathcal{D}_2(g_1)\mathcal{D}_2(g_2) v_2 \\
&= \mathcal{D}(g_1)(\mathcal{D}_1(g_2) v_1 \otimes \mathcal{D}_2(g_2) v_2) \\
&= \mathcal{D}(g_1)\mathcal{D}(g_2) v_1 \otimes v_2
\end{aligned} \tag{2.112}$$

so $\mathcal{D}(g_1 g_2) = \mathcal{D}(g_1)\mathcal{D}(g_2)$. Hence, this together with $\mathcal{D}(e) = 1$ implies that $\mathcal{D}(g)$ is invertible. So $\mathcal{D}(g)$ is a representation. ■

Note that if $\mathcal{D}_1$ is irreducible on V_1 and $\mathcal{D}_2$ is irreducible on V_2 then $\mathcal{D} = \mathcal{D}_1 \bigotimes \mathcal{D}_2$ is *not* generally irreducible on $V = V_1 \bigotimes V_2$. Indeed, we shall be particularly interested in decomposing $\mathcal{D}$ into irreducible components in several explicit examples.

2.15 Representations of Lie Algebras

Definition 29. *Let V be a finite dimensional vector space (over $\mathbb{R}$ or $\mathbb{C}$) and let $M(V)$ denote the space of linear transformations $V \to V$. Suppose that $\mathcal{L}(G)$ is the Lie algebra of a Lie group G. Then a representation of $\mathcal{L}(G)$ acting on V is a linear map $d : \mathcal{L}(G) \to M(V)$ satisfying*

$$d([X, Y]) = d(X)d(Y) - d(Y)d(X) \tag{2.113}$$

for all $X, Y \in \mathcal{L}(G)$. The dimension of the representation is the dimension of V.

Definition 30. *The trivial representation of $\mathcal{L}(G)$ on V is given by $d(X) = 0$ for all $X \in \mathcal{L}(G)$*

Definition 31. *If G is a matrix Lie group and hence $\mathcal{L}(G)$ is a matrix Lie algebra, then the tangent vectors can be regarded as matrices acting directly on n-component vectors. Then we define the fundamental representation of $\mathcal{L}(G)$ on V by $d(X) = X$*

There is a particularly natural representation associated with any Lie algebra.

Definition 32. *Let $\mathcal{L}(G)$ be a Lie algebra. Then the adjoint representation is a representation of $\mathcal{L}(G)$ over the vector space $\mathcal{L}(G)$, $\mathrm{ad} : \mathcal{L}(G) \to M(\mathcal{L}(G))$ defined by*

$$(\mathrm{ad}\ v)w = [v, w] \tag{2.114}$$

for $v, w \in \mathcal{L}(G)$.

It is clear from the above that $(\mathrm{ad}\ v)w$ is linear in w , hence $\mathrm{ad}\ v \in M(\mathcal{L}(G))$, and $\mathrm{ad}\ v$ is also linear in v.

Moreover, if $v_1, v_2, w \in \mathcal{L}(G)$ then

$$\begin{aligned}
(\mathrm{ad}\ [v_1, v_2])w &= [[v_1, v_2], w] \\
&= [v_1, [v_2, w]] - [v_2, [v_1, w]] \qquad \text{(using the Jacobi identity)} \\
&= (\mathrm{ad}\ v_1)[v_2, w] - (\mathrm{ad}\ v_2)[v_1, w] \\
&= (\mathrm{ad}\ v_1)(\mathrm{ad}\ v_2)w - (\mathrm{ad}\ v_2)(\mathrm{ad}\ v_1)w
\end{aligned} \tag{2.115}$$

so ad is indeed a representation.

2.16 The Baker-Campbell-Hausdorff (BCH) Formula

The BCH formula states that the product of two exponentials can be written as an exponential:

$$\exp(v)\exp(w) = \exp(v + w + \frac{1}{2}[v, w] + \frac{1}{12}[v, [v, w]] + \frac{1}{12}[[v, w], w] + ...) \tag{2.116}$$

where ... indicates terms of higher order in v and w. For simplicity we shall consider only matrix Lie groups, in which case the Lie algebra elements are square matrices.

To obtain the first few terms in this formula, consider $e^{tv}e^{tw}$ as a function of t and set

$$e^{Z(t)} = e^{tv}e^{tw} \tag{2.117}$$

At $t = 0$ we must have $e^{Z(0)} = \mathbb{I}$, which is solved by taking $Z(0) = 0$ (this solution is unique if we limit ourselves to the neighbourhood of the identity on which exp is invertible). Hence we can write $Z(t)$ as a power series

$$Z(t) = tP + \frac{1}{2}t^2 Q + O(t^3) \tag{2.118}$$

where we determine the matrices P and Q by expanding out (2.117) in powers of t:

$$\mathbb{I} + t(v + w) + \frac{1}{2}t^2(w^2 + v^2 + 2vw) + O(t^3) = \mathbb{I} + tP + \frac{1}{2}t^2(Q + P^2) + O(t^3) \tag{2.119}$$

from which we find $P = v + w$ and $Q = [v, w]$.

Proposition 13. *All higher order terms in the power series expansion of $Z(t)$ in the BCH formula depend only on sums of compositions of commutators on v and w.*

Proof

Suppose that $Z(y)$ is an arbitrary square matrix. Consider

$$f_1(x,y) = \frac{\partial}{\partial y}(e^{xZ(y)}) \tag{2.120}$$

and

$$f_2(x,y) = \int_{1-x}^{1} e^{(x-1+t)Z(y)} \frac{\partial Z(y)}{\partial y} e^{(1-t)Z(y)} dt \tag{2.121}$$

These both satisfy

$$\frac{\partial f_i}{\partial x} = \frac{\partial Z(y)}{\partial y} e^{xZ(y)} + Z(y) f_i(x,y) \tag{2.122}$$

and $f_i(0,y) = 0$ for $i = 1,2$. Hence $f_1(x,y) = f_2(x,y)$.
Now suppose that $Z(t)$ is the matrix appearing in the BCH formula, i.e.

$$e^{Z(t)} = e^{tv} e^{tw} \tag{2.123}$$

Then consider the identity $f_1(1,t) = f_2(1,t)$. This implies that

$$v + e^{tv} w e^{-tv} = \int_0^1 e^{yZ(t)} \frac{\partial Z(t)}{\partial t} e^{-yZ(t)} dy \tag{2.124}$$

Now consider the function

$$g(t) = e^{tv} w e^{-tv} \tag{2.125}$$

this satisfies $g(0) = Y$ and

$$\frac{d^n g}{dt^n} = e^{tv} (\text{ad } v)^n w e^{-tv} \tag{2.126}$$

hence the power series expansion of $g(t)$ is given by

$$e^{tv} w e^{-tv} = w + \sum_{n=1}^{\infty} \frac{t^n}{n!} (\text{ad } v)^n w \tag{2.127}$$

Applying this expression to both sides of (2.124) and performing the y-integral, one finds

$$v + w + \sum_{n=1}^{\infty} \frac{t^n}{n!} (\text{ad } v)^n w = \frac{dZ}{dt} + \sum_{n=1}^{\infty} \frac{1}{(n+1)!} (\text{ad } Z(t))^n \frac{dZ}{dt} \tag{2.128}$$

Then by expanding out $Z(t) = \sum_{n=1}^{\infty} Z_n t^n$ (as we know that $Z(0) = 0$), it follows by induction using the above equation that the Z_n can be written as sums of compositions of commutators on v and w. ■

Exercise: Suppose that $[v,w] = 0$. Show that $e^v e^w = e^{v+w}$.

Proposition 14. *Suppose that $\mathcal{D}$ is a representation of the matrix Lie group G acting on V. Then there is a representation d of $\mathcal{L}(G)$ also on V defined via*

$$d(v) = \frac{d}{dt}\big(\mathcal{D}(\exp(tv))\big)|_{t=0} \tag{2.129}$$

for $v \in \mathcal{L}(G)$.

<u>**Proof**</u>

It is convenient to expand out up to $O(t^3)$ by

$$\mathcal{D}(e^{tv}) = 1 + td(v) + t^2 h(v) + O(t^3) \tag{2.130}$$

Note that as $\mathcal{D}$ is a representation we must have $\mathcal{D}(e^{(t_1+t_2)v}) = \mathcal{D}(e^{t_1 v})\mathcal{D}(e^{t_2 v})$ Hence

$$1+(t_1+t_2)d(v)+(t_1+t_2)^2 h(v)+O(t_i^3) = (1+t_1 d(v)+t_1^2 h(v))(1+t_2 d(v)+t_2^2 h(v))+O(t_i^3) \tag{2.131}$$

and so on equating the $t_1 t_2$ coefficient we find $h(v) = \frac{1}{2}d(v)^2$.

Next consider for $v, w \in \mathcal{L}(G)$

$$\begin{aligned}\mathcal{D}(e^{-tv}e^{-tw}e^{tv}e^{tw}) &= (1 - td(v) + \frac{1}{2}t^2 d(v)^2)(1 - td(w) + \frac{1}{2}t^2 d(w)^2) \\ &\times (1 + td(v) + \frac{1}{2}t^2 d(v)^2)(1 + td(w) + \frac{1}{2}t^2 d(w)^2) + O(t^3) \\ &= 1 + t^2(d(v)d(w) - d(w)d(v)) + O(t^3)\end{aligned} \tag{2.132}$$

But using the BCH formula

$$\begin{aligned}e^{-tv}e^{-tw}e^{tv}e^{tw} &= e^{-t(v+w)+\frac{1}{2}t^2[v,w]+O(t^3)}e^{t(v+w)+\frac{1}{2}t^2[v,w]+O(t^3)} \\ &= e^{t^2[v,w]+O(t^3)}\end{aligned} \tag{2.133}$$

and so

$$\mathcal{D}(e^{-tv}e^{-tw}e^{tv}e^{tw}) = \mathcal{D}(e^{t^2[v,w]+O(t^3)}) = 1 + t^2 d([v,w]) + O(t^3) \tag{2.134}$$

Comparing (2.132) with (2.134) we find that

$$d([v,w]) = d(v)d(w) - d(w)d(v) \tag{2.135}$$

as required.

To show that d is linear, suppose $v, w \in \mathcal{L}(G)$ and α, β are constants. Then

$$\begin{aligned}\mathcal{D}(e^{t\alpha v}e^{t\beta w}) &= \mathcal{D}(e^{t\alpha v})\mathcal{D}(e^{t\beta w}) \\ &= (1 + t\alpha d(v) + O(t^2))(1 + t\beta d(w) + O(t^2)) \\ &= 1 + t(\alpha d(v) + \beta d(w)) + O(t^2)\end{aligned} \tag{2.136}$$

But by the BCH formula

$$\mathcal{D}(e^{t\alpha v}e^{t\beta w}) = \mathcal{D}(e^{t(\alpha v+\beta w)+O(t^2)}) = 1 + td(\alpha v + \beta w) + O(t^2) \tag{2.137}$$

Hence, comparing the $O(t)$ terms in (2.136) and (2.137) it follows that $d(\alpha v + \beta w) = \alpha d(v) + \beta d(w)$. ■

Proposition 15. *If G is a matrix Lie group, then the representation* Ad $: G \to GL(\mathcal{L}(G))$ *induces the representation* ad $: \mathcal{L}(G) \to M(\mathcal{L}(G))$.

Proof

If $v, w \in \mathcal{L}(G)$ then

$$\begin{aligned} \text{Ad }(e^{tv})w &= e^{tv}we^{-tv} \\ &= (1 + tv + O(t^2))w(1 - tv + O(t^2)) \\ &= w + t[v, w] + O(t^2) \\ &= (\mathbb{I} + t\text{ad } v + O(t^2))w \end{aligned} \tag{2.138}$$

and hence it follows that

$$\frac{d}{dt}(Ad(e^{tv}))|_{t=0} = \text{ad } v \tag{2.139}$$

as required. ■

We have seen that a representation $\mathcal{D}$ of the matrix Lie group G acting on V gives rise to a representation d of the Lie algebra $\mathcal{L}(G)$ on V. A partial converse is true.

Definition 33. *Suppose that G is a matrix Lie group. Let d denote a representation of $\mathcal{L}(G)$ on V. Then a representation $\mathcal{D}$ is induced locally on G via*

$$\mathcal{D}(g) = e^{d(v)} \tag{2.140}$$

for those $g \in G$ such that $g = e^v$.

Here we assume that the representation $d(v)$ is realized as a matrix linear transformation on V, so that the standard matrix exponentiation $e^{d(v)}$ may be taken. The representation $\mathcal{D}$ induced by d is generally not globally well-defined, but it is *locally* well-defined on the neighbourhood of the identity on which exp is invertible.

Proposition 16. *The map $\mathcal{D}$ given in (2.140) which is locally induced by the representation d of $\mathcal{L}(G)$ on V defines a representation.*

Proof

Clearly, $\mathcal{D}(g)$ defines a linear transformation on V.

As $\mathbb{I} = e^0$ it follows that $\mathcal{D}(e) = e^{d(0)} = e^0 = 1$ where $d(0) = 0$ follows from the linearity of d.

Also, suppose that g_1, g_2 have $g_1 = e^{v_1}$, $g_2 = e^{v_2}$. Then by the BCH formula we have

$$g_1 g_2 = e^{v_1+v_2+\frac{1}{2}[v_1,v_2]+\ldots} \tag{2.141}$$

where … denotes a sum of higher order nested commutators by Proposition 13. Hence

$$\begin{aligned}\mathcal{D}(g_1g_2) &= e^{d(v_1+v_2+\frac{1}{2}[v_1,v_2]+\dots)} \\ &= e^{d(v_1)+d(v_2)+\frac{1}{2}d([v_1,v_2])+d(\dots)} \quad \text{(by the linearity of } d\text{)} \\ &= e^{d(v_1)+d(v_2)+\frac{1}{2}[d(v_1),d(v_2)]+\dots} \quad \text{(using (2.113))} \\ &= e^{d(v_1)}e^{d(v_2)} \\ &= \mathcal{D}(g_1)\mathcal{D}(g_2)\end{aligned} \tag{2.142}$$

So $\mathcal{D}$ is at least locally a representation. Note that we have made use of the fact that all higher order terms in the BCH expansion can be written as sums of commutators, together with the property (2.113) of representations of $\mathcal{L}(G)$ in proceeding from the second to the third line of the above equation. ■

Proposition 17. *Suppose that G is a matrix Lie group. If $\mathcal{D}$ is a unitary representation of G on V then the induced representation d of $\mathcal{L}(G)$ on V is antihermitian.*

Conversely, suppose d is a antihermitian representation of $\mathcal{L}(G)$ on V, then the (locally) induced representation $\mathcal{D}$ of G on V is unitary.

Proof

First suppose that $\mathcal{D}$ is a unitary representation.

Recall that d satisfies $\mathcal{D}(e^{tX}) = \mathbb{I} + td(X) + O(t^2)$ for $t \in \mathbb{R}$ and $X \in \mathcal{L}(G)$.

As $\mathcal{D}$ is unitary it follows that $\mathcal{D}(e^{tX})\mathcal{D}(e^{tX})^\dagger = \mathbb{I}$.

Hence $(\mathbb{I} + td(X) + O(t^2))(\mathbb{I} + td(X) + O(t^2))^\dagger = \mathbb{I}$, so expanding out, the $O(t)$ terms imply $d(X) + d(X)^\dagger = 0$, i.e. d is antihermitian.

Conversely, suppose that d is an antihermitian representation of $\mathcal{L}(G)$ on V. Let $\mathcal{D}$ denote the (locally) induced representation of G on V.

Suppose that $g \in G$ is given by $g = e^X$ for $X \in \mathcal{L}(G)$. Then

$$\mathcal{D}(g) = e^{d(X)} \tag{2.143}$$

Then

$$\begin{aligned}\mathcal{D}(g)\mathcal{D}(g)^\dagger &= e^{d(X)}(e^{d(X)})^\dagger \\ &= e^{d(X)}e^{d(X)^\dagger} \\ &= e^{d(X)}e^{-d(X)} \\ &= 1\end{aligned} \tag{2.144}$$

Hence $\mathcal{D}(g)$ is unitary ■.

There are directly analogous definitions for irreducibility of representations of Lie algebras

Definition 34. *Suppose d is a representation of $\mathcal{L}(G)$ acting on V. A subspace $W \subset V$ is called an invariant subspace if $d(X)\omega \in W$ for all $X \in \mathcal{L}(G)$ and $\omega \in W$.*

Definition 35. *Suppose d is a representation of $\mathcal{L}(G)$ acting on V. Then d is reducible if there is an invariant subspace W of V with $W \neq 0$ and $W \neq V$. If the only invariant subspaces of V are 0 and V then d is called irreducible.*

Definition 36. *A representation d of $\mathcal{L}(G)$ is called totally reducible if there exists a direct sum decomposition of V into subspaces W_i, $V = W_1 \bigoplus W_2 \bigoplus ... \bigoplus W_k$ where the W_i are invariant subspaces with respect to d and d restricted to W_i is irreducible.*

Proposition 18. *Suppose that G is a matrix Lie group. If $\mathcal{D}$ is a representation of G on V with invariant subspace W, then W is an invariant subspace of the induced representation d of $\mathcal{L}(G)$ on V*

Conversely, suppose d is a representation of $\mathcal{L}(G)$ on V with invariant subspace W; then W is an invariant subspace of the (locally) induced representation $\mathcal{D}$ of G on V.

Proof

Suppose that $\mathcal{D}$ is a representation of G on V with invariant subspace W with respect to $\mathcal{D}$. Let d be the induced representation of $\mathcal{L}(G)$ on V. If $w \in W$ and $X \in \mathcal{L}(G)$ then

$$d(X)w = \frac{d}{dt}(\mathcal{D}(e^{tX}))_{t=0}w = \frac{d}{dt}(\mathcal{D}(e^{tX})w)_{t=0} \tag{2.145}$$

As $\mathcal{D}(e^{tX})w \in W$ for all $t \in \mathbb{R}$ it follows that $d(X)w \in W$.

Conversely, suppose that d is a representation of $\mathcal{L}(G)$ on V, and W is an invariant subspace of V with respect to d. Let $\mathcal{D}$ be the locally defined representation of G induced by d. Then if $g \in G$ is given by $g = e^X$ for some $X \in \mathcal{L}(G)$ then if $w \in W$,

$$\mathcal{D}(g)w = e^{d(X)}w = \sum_{n=0}^{\infty} \frac{1}{n!}d^n(X)w \tag{2.146}$$

However, as W is an invariant subspace of V with respect to d, it follows that $d^n(X)w \in W$ for all $n \in \mathbb{N}$. Hence $\mathcal{D}(g)w \in W$. ■

Note that in this proof we made implicit use of the closure of W.

There is also a natural concept of equivalent representations of Lie algebras.

Definition 37. *Suppose d is a representation of $\mathcal{L}(G)$ acting on V_1, and $B : V_1 \to V_2$ is an invertible linear transformation. Define $\tilde{d}(X) = Bd(X)B^{-1}$ for $X \in \mathcal{L}(G)$. Then d and $\tilde{d}$ are said to be equivalent representations.*

Exercise: Show that $\tilde{d}$ defined above is a representation of $\mathcal{L}(G)$. Also show that if $\mathcal{D}_1$, $\mathcal{D}_2$ are equivalent representations of G on vector spaces V_1 and V_2 then the corresponding induced representations d_1 and d_2 on V_1 and V_2 are equivalent; and conversely, if d_1 and d_2 are equivalent representations of $\mathcal{L}(G)$ on V_1 and V_2 then the locally defined induced representations $\mathcal{D}_1$ and $\mathcal{D}_2$ of G are equivalent.

Note that Schur's lemmas may be applied to representations of Lie algbras in exactly the same way as to representations of Lie groups.

Hence we have shown that there is (at least locally) a 1-1 correspondence between irreducible representations of the Lie group G and the Lie algebra $\mathcal{L}(G)$. This is useful, because it enables us to map the analysis of representations of G to those of $\mathcal{L}(G)$, and thus the problem reduces to one of linear algebra.

Proposition 19. *Suppose that d_1 and d_2 are representations of $\mathcal{L}(G)$ acting on V_1 and V_2 and let $V = V_1 \bigotimes V_2$. Define $d = d_1 \otimes 1 + 1 \otimes d_2$ as a linear map on V. Then d is a representation of $\mathcal{L}(G)$ acting on V.*

Proof

If $w_1, w_2 \in \mathcal{L}(V)$ and α, β are scalars and $v_1 \otimes v_2 \in V$ then

$$\begin{aligned} d(\alpha w_1 + \beta w_2) v_1 \otimes v_2 &= d_1(\alpha w_1 + \beta w_2) v_1 \otimes v_2 + v_1 \otimes d_2(\alpha w_1 + \beta w_2) v_2 \\ &= (\alpha d_1(w_1) + \beta d_1(w_2)) v_1 \otimes v_2 + v_1 \otimes (\alpha d_2(w_1) + \beta d_2(w_2)) v_2 \\ &= \alpha d_1(w_1) v_1 \otimes v_2 + \beta d_1(w_2) v_1 \otimes v_2 \\ &+ \alpha v_1 \otimes d_2(w_1) v_2 + \beta v_1 \otimes d_2(w_2) v_2 \\ &= \alpha (d_1(w_1) v_1 \otimes v_2 + v_1 \otimes d_2(w_1) v_2) \\ &+ \beta(\beta d_1(w_2) v_1 \otimes v_2 + v_1 \otimes d_2(w_2) v_2) \\ &= \alpha d(w_1) v_1 \otimes v_2 + \beta d(w_2) v_2 \otimes v_2 \end{aligned} \tag{2.147}$$

so d is linear on $\mathcal{L}(G)$. Also

$$\begin{aligned} d([w_1, w_2]) v_1 \otimes v_2 &= d_1([w_1, w_2]) v_1 \otimes v_2 + v_1 d_2([w_1, w_2]) v_2 \\ &= (d_1(w_1) d_1(w_2) - d_1(w_2) d_1(w_1)) v_1 \otimes v_2 \\ &+ v_1 \otimes (d_2(w_1) d_2(w_2) - d_2(w_2) d_2(w_1)) v_2 \\ &= d_1(w_1) d_1(w_2) v_1 \otimes v_2 - d_1(w_2) d_1(w_1) v_1 \otimes v_2 \\ &+ v_1 \otimes d_2(w_1) d_2(w_2) v_2 - v_1 \otimes d_2(w_2) d_2(w_1) v_2 \end{aligned} \tag{2.148}$$

Also note that

$$\begin{aligned} d(w_1) d(w_2) v_1 \otimes v_2 &= d(w_1)\big(d_1(w_2) v_1 \otimes v_2 + v_2 \otimes d_2(w_2) v_2\big) \\ &= d_1(w_1) d_1(w_2) v_1 \otimes v_2 + d_1(w_2) v_1 \otimes d_2(w_1) v_2 \\ &+ d_1(w_1) v_1 \otimes d_2(w_2) v_2 + v_1 \otimes d_2(w_1) d_2(w_2) v_2 \end{aligned} \tag{2.149}$$

where the sum of the second and third terms in this expression is symmetric in w_1 and w_2.

Hence

$$\begin{aligned} d(w_1) d(w_2) v_1 \otimes v_2 - d(w_2) d(w_1) v_1 \otimes v_2 &= d_1(w_1) d_1(w_2) v_1 \otimes v_2 - d_1(w_2) d_1(w_1) v_1 \otimes v_2 \\ &+ v_1 \otimes d_2(w_1) d_2(w_2) v_2 - v_1 \otimes d_2(w_2) d_2(w_1) v_2 \\ &= d([w_1, w_2]) v_1 \otimes v_2 \end{aligned} \tag{2.150}$$

as required ■.

Proposition 20. *Suppose that $\mathcal{D}_1$ and $\mathcal{D}_2$ are representations of matrix Lie group G on V_1 and V_2 with induced representations of $\mathcal{L}(G)$ on V_1 and V_2 denoted by d_1 and d_2. Let $\mathcal{D} = \mathcal{D}_1 \bigotimes \mathcal{D}_2$ denote the representation of G on the tensor product $V_1 \bigotimes V_2$. Then the corresponding induced representation of $\mathcal{L}(G)$ on $V_1 \bigotimes V_2$ is $d = d_1 \otimes 1 + 1 \otimes d_2$.*

Proof Suppose $w \in \mathcal{L}(G)$, then expanding out in powers of t;

$$\mathcal{D}(e^{tw}) v_1 \otimes v_2 = \mathcal{D}_1(e^{tw}) v_1 \otimes \mathcal{D}_2(e^{tw}) v_2$$

$$
\begin{aligned}
&= (v_1 + td_1(w)v_1 + O(t^2)) \otimes (v_2 + td_2(w)v_2 + O(t^2)) \\
&= v_1 \otimes v_2 + t\big(d_1(w)v_1 \otimes v_2 + v_1 \otimes d_2(w)v_2\big) + O(t^2) \\
&= v_1 \otimes v_2 + t(d_1 \otimes 1 + 1 \otimes d_2)v_1 \otimes v_2 + O(t^2)
\end{aligned} \tag{2.151}
$$

and hence from the $O(t)$ term we find the induced representation $d = d_1 \otimes 1 + 1 \otimes d_2$ as required ■.

2.17 The Killing Form and the Casimir Operator

Definition 38. *Suppose that G is a matrix Lie group with Lie algebra $\mathcal{L}(G)$. Then for $X \in \mathcal{L}(G)$,* ad X *can be realized as a matrix linear transformation on $\mathcal{L}(G)$. The Killing form κ is defined by*

$$\kappa(X,Y) = \mathrm{Tr}\ (\mathrm{ad}\ X \mathrm{ad}\ Y) \tag{2.152}$$

Suppose that T_a is a basis for $\mathcal{L}(G)$. Then from the Killing form one obtains a symmetric matrix

$$\kappa_{ab} = \kappa(T_a, T_b) \tag{2.153}$$

Denote the matrix elements of ad T_a by $(\mathrm{ad}\ T_a)_b{}^c$; then note that

$$
\begin{aligned}
(\mathrm{ad}\ T_a)T_b &= [T_a, T_b] = c_{ab}{}^c T_c \\
&= (\mathrm{ad}\ T_a)_b{}^c T_c
\end{aligned} \tag{2.154}
$$

Hence $(\mathrm{ad}\ T_a)_b{}^c = c_{ab}{}^c$. So it follows that

$$\kappa_{ab} = c_{ad}{}^e c_{be}{}^d \tag{2.155}$$

Lemma 4. *The Killing form is associative: $\kappa(X,[Y,Z]) = \kappa([X,Y],Z)$*
Proof

$$
\begin{aligned}
\kappa(X,[Y,Z]) &= \mathrm{Tr}\ (\mathrm{ad}\ X \mathrm{ad}\ [Y,Z]) \\
&= \mathrm{Tr}\ (\mathrm{ad}\ X(\mathrm{ad}\ Y \mathrm{ad}\ Z - \mathrm{ad}\ Z \mathrm{ad}\ Y)) \\
&= \mathrm{Tr}\ (\mathrm{ad}\ X \mathrm{ad}\ Y \mathrm{ad}\ Z) - \mathrm{Tr}\ (\mathrm{ad}\ X \mathrm{ad}\ Z \mathrm{ad}\ Y) \\
&= \mathrm{Tr}\ (\mathrm{ad}\ X \mathrm{ad}\ Y \mathrm{ad}\ Z) - \mathrm{Tr}\ (\mathrm{ad}\ Y \mathrm{ad}\ X \mathrm{ad}\ Z) \\
&= \mathrm{Tr}\ ((\mathrm{ad}\ X \mathrm{ad}\ Y - \mathrm{ad}\ Y \mathrm{ad}\ X)\mathrm{ad}\ Z) \\
&= \mathrm{Tr}\ (\mathrm{ad}\ [X,Y] \mathrm{ad}\ Z) \\
&= \kappa([X,Y],Z)
\end{aligned} \tag{2.156}
$$

as required. ■.

As κ_{ab} is symmetric, one can always choose an adapted basis for $\mathcal{L}(G)$ in which κ_{ab} is a diagonal matrix, and by rescaling the Lie algebra generators, the diagonal entries can be set to $+1$, 0 or -1.

Definition 39. *The Killing form is non-degenerate if κ_{ab} has no zero diagonal entries in the adapted basis. The Lie algebra $\mathcal{L}(G)$ is then called semi-simple. If all the diagonal entries are -1 then $\mathcal{L}(G)$ is said to be a compact Lie algebra.*

Lemma 5. *Suppose that $\mathcal{L}(G)$ is semi-simple. Define $c_{abc} = c_{ab}{}^d \kappa_{dc}$ (i.e. lower the last index of the structure constants with the Killing form). Then c_{abc} is totally antisymmetric in a, b, c.*

Proof

Note that

$$\kappa(T_a, [T_b, T_c]) = c_{bc}{}^d \kappa(T_a, T_d) = c_{bc}{}^d \kappa_{ad} = c_{bca} \tag{2.157}$$

and

$$\kappa([T_a, T_b], T_c) = c_{ab}{}^d \kappa(T_d, T_c) = c_{ab}{}^d \kappa_{dc} = c_{abc} = -c_{bac} \tag{2.158}$$

But by associativity of κ, (2.157) and (2.158) are equal. Hence $c_{bca} = -c_{bac}$. As c_{abc} is skew symmetric in both the first two and the last two indices, it follows that c_{abc} is totally antisymmetric. ■

Definition 40. *Suppose that $\mathcal{L}(G)$ is a Lie algebra with non-degenerate Killing form, and d is a representation of $\mathcal{L}(G)$ on V. The Casimir operator C of $\mathcal{L}(G)$ is defined by*

$$C = -\sum_{a,b} (\kappa^{-1})^{ab} d(T_a) d(T_b) \tag{2.159}$$

where κ^{-1} denotes the inverse of the Killing form.

Proposition 21. *The Casimir operator commutes with $d(X)$ for all $X \in \mathcal{L}(G)$*

Proof

It suffices to show that $[C, d(T_a)] = 0$ for all T_a.

Note that

$$\begin{aligned}
[d(T_a), C] &= -\sum_{b,c} (\kappa^{-1})^{bc} \big([d(T_a), d(T_b) d(T_c)] \big) \\
&= -\sum_{b,c} (\kappa^{-1})^{bc} \big([d(T_a), d(T_b)] d(T_c) + d(T_b) [d(T_a), d(T_c)] \big) \\
&= -\sum_{b,c} (\kappa^{-1})^{bc} \big(d[T_a, T_b] d(T_c) + d(T_b) d[T_a, T_c] \big) \\
&= -\sum_{b,c} (\kappa^{-1})^{bc} \big(c_{ab}{}^{\ell} d(T_\ell) d(T_c) + c_{ac}{}^{\ell} d(T_b) d(T_\ell) \big) \\
&= -c_a{}^{c\ell} d(T_\ell) d(T_c) - c_a{}^{c\ell} d(T_c) d(T_\ell) \\
&= 0
\end{aligned} \tag{2.160}$$

where we have defined

$$c_a{}^{bc} = \sum_\ell (\kappa^{-1})^{b\ell} c_{a\ell}{}^c \tag{2.161}$$

and we note that $c_a{}^{bc}$ is antisymmetric in b, c. ■

Note that if d is irreducible then by Schur's second lemma, C must be a scalar multiple of the identity.

If $\mathcal{L}(G)$ is compact, then working in the adapted basis, C takes a particularly simple form

$$C = \sum_a d(T_a) d(T_a) \tag{2.162}$$

3. SU(2) and Isospin

3.1 Lie Algebras of $SO(3)$ and $SU(2)$

We have shown that the Lie algebra of $SO(3)$ consists of the 3×3 antisymmetric real matrices.

A basis for $\mathcal{L}(SO(3))$ is given by

$$T_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad T_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad T_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.1}$$

Exercise: Show that

$$[T_a, T_b] = \epsilon_{abc} T_c \tag{3.2}$$

where ϵ_{abc} is totally antisymmetric and $\epsilon_{123} = 1$.

Next consider $SU(n)$. Let $M(t)$ be a smooth curve in $SU(n)$ with $M(0) = \mathbb{I}$. Then $M(t)$ must satisfy

$$\det M(t) = 1 \qquad M(t)M(t)^\dagger = \mathbb{I} \tag{3.3}$$

Differentiating these constraints we obtain

$$\mathrm{Tr}\left(M(t)^{-1}\frac{dM(t)}{dt}\right) = 0 \qquad \frac{dM(t)}{dt}M(t)^\dagger + M(t)\frac{dM(t)^\dagger}{dt} = 0 \tag{3.4}$$

Setting $t = 0$ we find

$$\mathrm{Tr}\ m = 0 \qquad m + m^\dagger = 0 \tag{3.5}$$

where $m = \frac{dM(t)}{dt}|_{t=0}$. Hence $\mathcal{L}(SU(n))$ consists of the traceless antihermitian matrices.

Exercise: Verify that if X and Y are traceless antihermitian square matrices then so is $[X, Y]$.

It is convenient to make use of the Pauli matrices σ_a defined by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.6}$$

which satisfy $\sigma_a \sigma_b = \delta_{ab}\mathbb{I} + i\epsilon_{abc}\sigma_c$.

Then a basis of traceless antihermitian 2×2 matrices is given by taking $T_a = -\frac{i}{2}\sigma_a$.

It follows that

$$[T_a, T_b] = \epsilon_{abc} T_c \tag{3.7}$$

Comparing (3.2) and (3.7) it is clear that $SO(3)$ and $SU(2)$ have the same Lie algebra. We might therefore expect $SO(3)$ and $SU(2)$ to be similar, at least near to the identity. We shall see that this is true.

Exercise: Using this basis, show that the Lie algebra $\mathcal{L}(SU(2))$ is of compact type.

3.2 Relationship between $SO(3)$ and $SU(2)$

Proposition 22. *The manifold $SU(2)$ can be identified with S^3.*

Proof

Consider

$$U = \begin{pmatrix} \alpha & \mu \\ \beta & \nu \end{pmatrix} \in SU(2) \tag{3.8}$$

for α, β, μ, $\nu \in \mathbb{C}$.

Then $UU^\dagger = \mathbb{I}$ implies that $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ be orthogonal to $\begin{pmatrix} \mu \\ \nu \end{pmatrix}$ with respect to the standard inner product on $\mathbb{C}^2$. As the orthogonal complement to $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ in $\mathbb{C}^2$ is a 1-dimensional complex vector space spanned by $\begin{pmatrix} -\bar{\beta} \\ \bar{\alpha} \end{pmatrix}$ it follows that $\mu = -\sigma\bar{\beta}$, $\nu = \sigma\bar{\alpha}$ for $\sigma \in \mathbb{C}$.

We also require that $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ and $\begin{pmatrix} \mu \\ \nu \end{pmatrix}$ have unit norm, which fixes

$$|\sigma|^2 = |\alpha|^2 + |\beta|^2 = 1 \tag{3.9}$$

Finally, the constraint $\det U = 1$ fixes $\sigma = 1$.

Hence we have shown that

$$U = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} \tag{3.10}$$

where $\alpha\bar{\alpha} + \beta\bar{\beta} = 1$. Such U is automatically an element of $SU(2)$. Hence this is the form of the most generic element of $SU(2)$.

Set $\alpha = y_0 + iy_3$, $\beta = -y_2 + iy_1$ for $y_0, y_1, y_2, y_3 \in \mathbb{R}$. Then it is straightforward to see that

$$U = y_0\mathbb{I} + iy_n\sigma_n \tag{3.11}$$

and $\alpha\bar{\alpha} + \beta\bar{\beta} = 1$ is equivalent to $y_0^2 + y_1^2 + y_2^2 + y_3^2 = 1$, i.e. $y \in S^3$.

This establishes a smooth invertible map between $SU(2)$ and S^3. ■

Note that this map provides an explicit way of seeing that $SU(2)$ is connected, because S^3 is connected.

Proposition 23. *There is a $2-1$ correspondence $R : SU(2) \to SO(3)$ between $SU(2)$ and $SO(3)$. The map R is a group homomorphism.*

Proof Suppose $U \in SU(2)$. Define a 3×3 matrix $R(U)$ via

$$R(U)_{mn} = \frac{1}{2}\mathrm{Tr}\,\left(\sigma_m U \sigma_n U^\dagger\right) \tag{3.12}$$

By writing $U = y_0\mathbb{I} + iy_m\sigma_m$ for $y_0, y_m \in \mathbb{R}$ satisfying $y_0^2 + y_p y_p = 1$, it is straightforward to show that

$$R_{mn} = (y_0^2 - y_p y_p)\delta_{mn} + 2\epsilon_{mnq}y_0 y_q + 2y_m y_n \tag{3.13}$$

(here we have written $R_{mn} = R(U)_{mn}$).

It is clear that if $y_p = 0$ for $p = 1, 2, 3$ so that $U = \pm\mathbb{I}$, then $R = \mathbb{I}$, so $R \in SO(3)$.

More generally, suppose that $y_p y_p \neq 0$. Then one can set $y_0 = \cos\alpha$, $y_p = \sin\alpha z_p$ for $0 < \alpha < 2\pi$ and $\alpha \neq \pi$. Then the constraint $y_0^2 + y_p y_p = 1$ implies that $z_p z_p = 1$, i.e $\underline{z}$ is a unit vector in $\mathbb{R}^3$. The expression (3.13) can be rewritten as

$$R_{mn} = \cos 2\alpha \delta_{mn} + \sin 2\alpha \epsilon_{mnq} z_q + (1 - \cos 2\alpha) z_m z_n \tag{3.14}$$

It is then apparent that

$$R_{mn} z_n = z_m \tag{3.15}$$

and if $\underline{x}$ is orthogonal to $\underline{z}$ then

$$R_{mn} x_n = \cos 2\alpha x_m + \sin 2\alpha \epsilon_{mnq} x_n z_q \tag{3.16}$$

The transformation R therefore corresponds to a rotation of angle 2α in the plane with unit normal vector $\underline{z}$.

It is clear that any non-trivial rotation in $SO(3)$ can be written in this fashion. However, the correspondence is not $1-1$. To see this explicitly, suppose that two rotations corresponding to $y \in S^3$ and $u \in S^3$ are equal. Then

$$(y_0^2 - y_p y_p)\delta_{mn} + 2\epsilon_{mnq} y_0 y_q + 2 y_m y_n = (u_0^2 - u_p u_p)\delta_{mn} + 2\epsilon_{mnq} u_0 u_q + 2 u_m u_n \tag{3.17}$$

where $y_0^2 + y_p y_p = u_0^2 + u_p u_p = 1$.

From the antisymmetric portion of this matrix we find $y_0 y_p = u_0 u_p$.

From the diagonal elements with $n = m$, we see that

$$y_0^2 - y_p y_p + 2(y_m)^2 = u_0^2 - u_p u_p + 2(u_m)^2 \tag{3.18}$$

where p is summed over but m is fixed. Summing over m we find

$$3y_0^2 - y_p y_p = 3u_0^2 - u_p u_p \tag{3.19}$$

where p is summed over; which together with $y_0^2 + y_p y_p = u_0^2 + u_p u_p = 1$ implies that $y_0^2 = u_0^2$ and $y_p y_p = u_p u_p$ (sum over p). Substituting back into (3.18) we find $y_m^2 = u_m^2$ for *each* $m = 1, 2, 3$.

Suppose first that $y_0 \neq 0$, then $u_0 \neq 0$, and it follows that $y_0 = \pm u_0$ and $y_p = \pm u_p$ for each $p = 1, 2, 3$ (with the same sign throughout).

Next, suppose $y_0 = 0$. Then $u_0 = 0$ also, and $y_m y_n = u_m u_n$ for each $m, n = 1, 2, 3$. Contracting with y_n we get $(y_n y_n) y_m = (y_n u_n) u_m$ for $m = 1, 2, 3$. As $y_n y_n = 1$, this implies that $y_m = \lambda u_m$ for $m = 1, 2, 3$ where λ is constant. Hence, $(1 - \lambda^2) u_m u_n = 0$. Contracting over m and n then gives $1 - \lambda^2 = 0$, so $\lambda = \pm 1$. Therefore $y_p = \pm u_p$ for $p = 1, 2, 3$ (with the same sign throughout).

Hence, we have shown that each $R \in SO(3)$ corresponds to two elements U and $-U \in SU(2)$. These two elements correspond to antipodal points $\pm y \in S^3$. This establishes the correspondence.

It remains to check that $R(U_1U_2) = R(U_1)R(U_2)$ for $U_1, U_2 \in SU(2)$. Note that on writing

$$U_1 = y_0\mathbb{I}_2 + iy_n\sigma_n, \quad U_2 = w_0 + iw_n\sigma_n \tag{3.20}$$

for $y_0, y_p, w_0, w_p \in \mathbb{R}$ satisfying $y_0^2 + y_py_p = w_0^2 + w_pw_p = 1$ then

$$U_1U_2 = u_0\mathbb{I}_2 + iu_n\sigma_n \tag{3.21}$$

where $u_0 = y_0w_0 - y_pw_p$ and $u_m = y_0w_m + w_0y_m - \epsilon_{mpq}y_pw_q$ satisfy $u_0^2 + u_pu_p = 1$. It then suffices to evaluate directly

$$R(U_1U_2)_{mn} = (u_0^2 - u_pu_p)\delta_{mn} + 2\epsilon_{mnq}u_0u_q + 2u_mu_n \tag{3.22}$$

and compare this with

$$R(U_1)_{mp}R(U_2)_{pn} = \left[(y_0^2 - y_\ell y_\ell)\delta_{mp} + 2\epsilon_{mpq}y_0y_q + 2y_my_p\right]\left[(w_0^2 - w_rw_r)\delta_{pn} + 2\epsilon_{pnr}w_0w_r + 2w_pw_n\right] \tag{3.23}$$

On expanding out these two expressions in terms of y and w it becomes an exercise in algebra to show that $R(U_1U_2)_{mn} = R(U_1)_{mp}R(U_2)_{pn}$ as required. ■

Exercise: Verify the identity $R(U_1U_2)_{mn} = R(U_1)_{mp}R(U_2)_{pn}$.

It is conventional to write $SU(2) = S^3$ and $SO(3) = S^3/\mathbb{Z}_2$, where $S^3/\mathbb{Z}_2$ is the 3-sphere with antipodal points identified. $SU(2)$ is called the double cover of $SO(3)$; and there is an isomorphism $SO(3) \cong SU(2)/\mathbb{Z}_2$.

It can be shown (using topological methods outside the scope of this course) that $SU(2)$ and $SO(3)$ are *not* homeomorphic. This is because $SU(2)$ and $SO(3)$ have different fundamental groups π_1. In particular, as $SU(2) \cong S^3$, and S^3 is simply connected, it follows that $\pi_1(SU(2))$ is trivial. However, it can be shown that $\pi_1(SO(3)) = \mathbb{Z}_2$.

3.3 Irreducible Representations of $SU(2)$

The standard basis of $\mathcal{L}(SU(2))$ which we have been using is $T_a = -\frac{i}{2}\sigma_a$. Suppose that d is a finite-dimensional irreducible representation of $\mathcal{L}(SU(2))$ on V, where V is a vector space over $\mathbb{C}$.

Define

$$J_3 = id(T_3), \qquad J_\pm = \frac{i}{\sqrt{2}}(d(T_1) \pm id(T_2)) \tag{3.24}$$

Then

$$[J_3, J_\pm] = \pm J_\pm, \qquad [J_+, J_-] = J_3 \tag{3.25}$$

As V is a complex vector space, there exists an eigenstate $|\phi\rangle$ of J_3 with some eigenvalue λ, and using (3.25) it follows that

$$J_3 J_\pm |\phi\rangle = (\lambda \pm 1) J_\pm |\phi\rangle \tag{3.26}$$

and so by simple induction

$$J_3 (J_\pm)^n |\phi\rangle = (\lambda \pm n)(J_\pm)^n |\phi\rangle \tag{3.27}$$

for non-negative integer n; i.e. $J_\pm$ are the standard raising and lowering operators. The eigenvalues of J_3 are called weights. We have therefore shown that $(J_\pm)^n |\phi\rangle$ either vanishes or is a J_3 eigenstate with eigenvalue $\lambda \pm n$.

Consider the states $(J_+)^n |\phi\rangle$. If non-vanishing, these states are linearly independent (as they have different J_3 eigenvalues). Hence, as V is finite dimensional, there exists a J_3 eigenstate $(J_+)^n |\phi\rangle$ which we will refer to as $|j, j\rangle$ with eigenvalue j such that $J_+ |j, j\rangle = 0$.

Definition 41. *j is called the highest weight of the representation. In the context of particle physics, it is called the spin.*

Note that by acting on $|j, j\rangle$ with J_- other distinct eigenstates of J_3 are obtained. As we are interested in finite dimensional representations, it follows that $(J_-)^N |j, j\rangle = 0$ for some positive integer N (otherwise one could just keep going and the representation would be infinite dimensional). Let $k + 1$ be the smallest positive integer for which this happens, and set $|\psi_k\rangle = (J_-)^k |j, j\rangle$, so, by construction, $J_- |\psi_k\rangle = 0$.

Define for $\ell = 0, \ldots, k$

$$|\psi_\ell\rangle = (J_-)^\ell |j, j\rangle \tag{3.28}$$

Then $|\psi_\ell\rangle$ are (non-vanishing) $j - \ell$ eigenvectors of J_3.

Note that

$$\begin{aligned} J_+ |\psi_\ell\rangle &= J_+ (J_-)^\ell |j, j\rangle \\ &= (J_+ J_-)(J_-)^{\ell-1} |j, j\rangle \\ &= \big([J_+, J_-] + J_- J_+\big)(J_-)^{\ell-1} |j, j\rangle \\ &= \big(J_3 + J_- J_+\big)(J_-)^{\ell-1} |j, j\rangle \\ &= (j - (\ell - 1))(J_-)^{\ell-1} |j, j\rangle + J_- J_+ (J_-)^{\ell-1} |j, j\rangle \end{aligned} \tag{3.29}$$

Repeating this process, one finds by induction

$$\begin{aligned} J_+ |\psi_\ell\rangle &= \big(j - (\ell - 1) + j - (\ell - 2) + \cdots + j - 1 + j\big) |\psi_{\ell-1}\rangle \\ &= \ell(j - \frac{1}{2}(\ell - 1)) |\psi_{\ell-1}\rangle \end{aligned} \tag{3.30}$$

In order to constrain j recall that $J_- |\psi_k\rangle = 0$, so

$$\begin{aligned} 0 &= J_+ J_- |\psi_k\rangle \\ &= \big([J_+, J_-] + J_- J_+\big) |\psi_k\rangle \end{aligned}$$

$$
\begin{aligned}
&= \big(J_3 + J_- J_+\big) |\psi_k\rangle \\
&= (j-k) |\psi_k\rangle + J_-\big(k(j - \frac{1}{2}(k-1))\big) |\psi_{k-1}\rangle \\
&= \big(j - k + k(j - \frac{1}{2}(k-1))\big) |\psi_k\rangle \\
&= \frac{1}{2}(k+1)(2j-k) |\psi_k\rangle
\end{aligned}
\tag{3.31}
$$

As k is non-negative, it follows that $k = 2j$

Using this it is straightforward to prove the following

Proposition 24. $V = \mathrm{span}\{J_-^k |j,j\rangle, J_-^{k-1} |j,j\rangle, \ldots, |j,j\rangle\}$ *and the highest weight state is unique.*

<u>Proof</u>

Consider the vector space V' spanned by $|\psi_i\rangle$ for $i = 0, \ldots, k$. This is an invariant subspace of V with respect to the representation d. As the representation is irreducible on V it follows that $V = V'$. In particular, J_3 is diagonalizable on V and each eigenspace is 1-dimensional.

To prove uniqueness suppose that $|\phi\rangle \in V$ satsfies $J_+ |\phi\rangle = 0$. As $|\phi\rangle \in V$ it follows that we can write

$$
|\phi\rangle = \sum_{i=0}^{2j} a_i (J_-)^i |j,j\rangle \tag{3.32}
$$

for constants a_i. Applying $(J_+)^{2j}$ to both sides of this equation implies $a_{2j} = 0$. Then, applying $(J_+)^{2j-1}$ implies $a_{2j-1} = 0$. Continuing in this way, we obtain $a_1 = a_2 = \cdots = a_{2j-1} = a_{2j} = 0$, and so $|\phi\rangle = a_0 |j,j\rangle$. So the highest weight state in an irreducible representation of $\mathcal{L}(SU(2))$ is unique (up to scaling). $\blacksquare$.

Hence, we have shown that j is half (non-negative) integer, and the representation is $2j+1$-dimensional. The irreducible representations are therefore completely characterized by the value of the weight j.

It is possible to go further, and prove the following theorem (the proof given is that presented in [Samelson]):

Theorem 1. ***NON-EXAMINABLE** Suppose that d is a representation of $\mathcal{L}(SU(2))$ on a complex vector space V. Then V can be decomposed as $V = V_1 \oplus \cdots \oplus V_p$ where V_i are invariant subspaces of V such that d restricted to V_i is irreducible.*

<u>Proof</u> Given in Appendix A.

Note that we have *not* as yet assumed that the representation originates from a unitary representation of the Lie group. However, in order to compute normalized states, it will be convenient to assume this; so in particular, the $d(T_a)$ are antihermitian and hence

$$
J_3^\dagger = J_3, \qquad J_\pm^\dagger = J_\mp \tag{3.33}
$$

We will assume that the highest weight state $|j,j\rangle$ is normalized to 1.

<u>Exercise:</u>

Show using (3.30) that $\langle\psi_\ell|\,\psi_\ell\rangle = \frac{\ell}{2}(2j-\ell+1)\,\langle\psi_{\ell-1}|\,\psi_{\ell-1}\rangle$, and hence that

$$N_\ell \equiv \langle\psi_\ell|\,\psi_\ell\rangle = \frac{(2j)!\ell!}{2^\ell(2j-\ell)!} \tag{3.34}$$

It is conventional to define normalized states $|j,m\rangle$ for $m=-j,\dots,j$ via

$$|j,m\rangle = \frac{1}{\sqrt{N_{j-m}}}\,|\psi_{j-m}\rangle \tag{3.35}$$

the first label denotes the highest weight value j, the m label denotes the eigenstate of J_3, $J_3\,|j,m\rangle = m\,|j,m\rangle$. These satisfy (check!)

$$\begin{aligned} J_-\,|j,m\rangle &= \frac{1}{\sqrt{2}}\sqrt{(j+m)(j-m+1)}\,|j,m-1\rangle \\ J_+\,|j,m-1\rangle &= \frac{1}{\sqrt{2}}\sqrt{(j+m)(j-m+1)}\,|j,m\rangle \end{aligned} \tag{3.36}$$

<u>Exercise:</u> Show that the Casimir operator C is given by $C = -\frac{1}{2}(J_+J_- + J_-J_+ + (J_3)^2)$ and satisfies

$$C\,|j,m\rangle = -\frac{1}{2}j(j+1)\,|j,m\rangle \tag{3.37}$$

3.3.1 Examples of Low Dimensional Irreducible Representations

It is useful to consider several low-dimensional representations.

i) $j=\frac{1}{2}$. A normalized basis is given by $|\frac{1}{2},\frac{1}{2}\rangle$ and $|\frac{1}{2},-\frac{1}{2}\rangle$, with

$$\begin{aligned} J_+\,|\frac{1}{2},\frac{1}{2}\rangle &= 0 \\ J_+\,|\frac{1}{2},-\frac{1}{2}\rangle &= \frac{1}{\sqrt{2}}\,|\frac{1}{2},\frac{1}{2}\rangle \end{aligned} \tag{3.38}$$

and

$$\begin{aligned} J_-\,|\frac{1}{2},\frac{1}{2}\rangle &= \frac{1}{\sqrt{2}}\,|\frac{1}{2},-\frac{1}{2}\rangle \\ J_-\,|\frac{1}{2},-\frac{1}{2}\rangle &= 0 \end{aligned} \tag{3.39}$$

ii) $j=1$. A normalized basis of states is $|1,1\rangle$, $|1,0\rangle$ and $|1,-1\rangle$ with

$$\begin{aligned} J_+\,|1,1\rangle &= 0 \\ J_+\,|1,0\rangle &= |1,1\rangle \\ J_+\,|1,-1\rangle &= |1,0\rangle \end{aligned} \tag{3.40}$$

and

$$
\begin{aligned}
J_- |1,1\rangle &= |1,0\rangle \\
J_- |1,0\rangle &= |1,-1\rangle \\
J_- |1,-1\rangle &= 0
\end{aligned} \tag{3.41}
$$

iii) $j = \frac{3}{2}$. A normalized basis of states is $|\frac{3}{2},\frac{3}{2}\rangle$, $|\frac{3}{2},\frac{1}{2}\rangle$, $|\frac{3}{2},-\frac{1}{2}\rangle$ and $|\frac{3}{2},-\frac{3}{2}\rangle$. with

$$
\begin{aligned}
J_+ |\frac{3}{2},\frac{3}{2}\rangle &= 0 \\
J_+ |\frac{3}{2},\frac{1}{2}\rangle &= \sqrt{\frac{3}{2}}\,|\frac{3}{2},\frac{3}{2}\rangle \\
J_+ |\frac{3}{2},-\frac{1}{2}\rangle &= \sqrt{2}\,|\frac{3}{2},\frac{1}{2}\rangle \\
J_+ |\frac{3}{2},-\frac{3}{2}\rangle &= \sqrt{\frac{3}{2}}\,|\frac{3}{2},-\frac{1}{2}\rangle
\end{aligned} \tag{3.42}
$$

and

$$
\begin{aligned}
J_- |\frac{3}{2},\frac{3}{2}\rangle &= \sqrt{\frac{3}{2}}\,|\frac{3}{2},\frac{1}{2}\rangle \\
J_- |\frac{3}{2},\frac{1}{2}\rangle &= \sqrt{2}\,|\frac{3}{2},-\frac{1}{2}\rangle \\
J_- |\frac{3}{2},-\frac{1}{2}\rangle &= \sqrt{\frac{3}{2}}\,|\frac{3}{2},-\frac{3}{2}\rangle \\
J_- |\frac{3}{2},-\frac{3}{2}\rangle &= 0
\end{aligned} \tag{3.43}
$$

3.4 Tensor Product Representations

Suppose that d_1 and d_2 are two irreducible representations of $\mathcal{L}(SU(2))$ over vector spaces $V(1)$, $V(2)$. Let $V = V(1) \bigotimes V(2)$ be the tensor product, and $d = d_1 \otimes 1 + 1 \otimes d_2$ be the representation on V.

We wish to decompose V into irreducible representations of d (i.e. invariant subspaces of V on which the restriction of d is irreducible).

Denote the states of $V(1)$ by $|j_1, m\rangle$ for $m = -j_1, \dots, j_1$ and those of $V(2)$ by $|j_2, n\rangle$ for $n = -j_2, \dots, j_2$, where j_1 and j_2 are the highest weights of $V(1)$ and $V(2)$ respectively. Note that $|j_1, m\rangle \otimes |j_2, n\rangle$ for $m = -j_1, \dots, j_1$ and $n = -j_2, \dots, j_2$ provides a basis for V.

Set

$$
J_3(1) = id_1(T_3), \qquad J_\pm(1) = \frac{i}{\sqrt{2}}(d_1(T_1) \pm id_1(T_2)) \tag{3.44}
$$

and

$$
J_3(2) = id_2(T_3), \qquad J_\pm(2) = \frac{i}{\sqrt{2}}(d_2(T_1) \pm id_2(T_2)) \tag{3.45}
$$

with

$$J_3 = J_3(1) \otimes 1 + 1 \otimes J_3(2), \qquad J_\pm = J_\pm(1) \otimes 1 + 1 \otimes J_\pm(2) \tag{3.46}$$

Exercise: Check that $J_\pm$ and J_3 satisfy $J_3^\dagger = J_3$, $J_\pm^\dagger = J_\mp$ and

$$[J_3, J_\pm] = \pm J_\pm, \qquad [J_+, J_-] = J_3 \tag{3.47}$$

In order to construct the decomposition of V into irreducible representations, first note that if $|\psi(1)\rangle \in V(1)$ is a state of weight m_1 with respect to $J_3(1)$, and $|\psi(2)\rangle \in V(2)$ is a state of weight m_2 with respect to $J_3(2)$ then $|\psi(1)\rangle \otimes |\psi(2)\rangle \in V$ is a state of weight $m_1 + m_2$ with respect to J_3, i.e. weights add in the tensor product representation.

Using this, it is possible to compute the degeneracy of certain weight states in the tensor product representation. In particular, the maximum possible weight must be $j_1 + j_2$ which corresponds to $|j_1, j_1\rangle \otimes |j_1, j_2\rangle$.

Consider the weight $j_1 + j_2 - k$ for $k > 0$. In general, $j_1 + j_2 - k$ can be written as a sum of two integers $m_1 + m_2$, for $m_1 \in \{-j_1, ..., j_1\}$ and $m_2 \in \{-j_2, ..., j_2\}$ in $k + 1$ ways:

$$\begin{aligned} j_1 + j_2 - k &= (j_1 - k) + j_2 \\ &= (j_1 - k + 1) + (j_2 - 1) \\ &\dots \\ &= (j_1 - 1) + (j_2 - k + 1) \\ &= j_1 + (j_2 - k) \end{aligned} \tag{3.48}$$

provided that $j_1 - k \geq -j_1$ and $j_2 - k \geq -j_2$, or equivalently

$$k \leq 2\min(j_1, j_2) = j_1 + j_2 - |j_1 - j_2| \tag{3.49}$$

Consider the state of weight $j_1 + j_2$. There is only one such state, and there is no state of higher weight. Hence it must be a highest weight state of a subspace of V on which the tensor product representation is irreducible. This irreducible subspace has dimension $2(j_1 + j_2) + 1$, and is denoted by $V_{j_1+j_2}$ Write $V = V' \oplus V_{j_1+j_2}$.

Next consider the states in V'. The highest possible weight is $j_1 + j_2 - 1$. In V there were two linearly independent states of this weight, however one of these lies in $V_{j_1+j_2}$, and does not lie in V'. The remaining state of weight $j_1 + j_2 - 1$ is of highest weight in V' and hence must be a highest weight state of a subspace (of dimension $2(j_1 + j_2) - 1$) of V', on which the tensor product representation is irreducible. Denote this subspace by $V_{j_1+j_2-1}$. Note that by construction $V_{j_1+j_2} \cap V_{j_1+j_2-1} = \{0\}$.

One can continue inductively in this fashion: for each $j = |j_1 - j_2|, |j_1 - j_2| + 1, \dots\ j_1 + j_2$ there is a subspace V_j (of dimension $2j+1$) on which the tensor product representation is irreducible, with highest weight j, and $V_j \cap V_k = \{0\}$ if $j \neq k$. In fact these subspaces account for all of the elements in V. To see this, compute the dimension of $V_{|j_1-j_2|} \bigoplus \cdots \bigoplus V_{j_1+j_2}$:

We find

$$\dim V_{|j_1-j_2|} \bigoplus \cdots \bigoplus V_{j_1+j_2} = \sum_{j=|j_1-j_2|}^{j_1+j_2} (2j+1)$$

$$
\begin{aligned}
&= \sum_{n=0}^{j_1+j_2-|j_1-j_2|} (2(|j_1-j_2|+n)+1) \\
&= (1+2j_1)(1+2j_2) \\
&= \dim\, V
\end{aligned} \tag{3.50}
$$

Hence we have decomposed $V = V_{|j_1-j_2|} \bigoplus \cdots \bigoplus V_{j_1+j_2}$ into irreducible subspaces V_j where V_j has highest weight j and is of dimension $2j+1$.

3.4.1 Examples of Tensor Product Decompositions

We shall consider two simple spin combinations which are of physical interest.

Firstly, take the spin $1/2 \otimes$ spin $1/2$ composite system. As $j_1 = j_2 = \frac{1}{2}$ there are two possible values for the composite spin, $j = 1$ or $j = 0$, and the tensor product space is 4-dimensional.

For the $j = 1$ states, the highest weight state can be written as

$$
|1,1\rangle = |\frac{1}{2},\frac{1}{2}\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle \tag{3.51}
$$

Applying J_- to both sides we find

$$
|1,0\rangle = \frac{1}{\sqrt{2}}\big(\,|\frac{1}{2},-\frac{1}{2}\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle + |\frac{1}{2},\frac{1}{2}\rangle \otimes |\frac{1}{2},-\frac{1}{2}\rangle\,\big) \tag{3.52}
$$

and applying J_- once more

$$
|1,-1\rangle = |\frac{1}{2},-\frac{1}{2}\rangle \otimes |\frac{1}{2},-\frac{1}{2}\rangle \tag{3.53}
$$

The remaining possible spin is $j = 0$. This has only one state, which must have the form

$$
|0,0\rangle = c_0\,|\frac{1}{2},\frac{1}{2}\rangle \otimes |\frac{1}{2},-\frac{1}{2}\rangle + c_1\,|\frac{1}{2},-\frac{1}{2}\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle \tag{3.54}
$$

for constants c_0, c_1 to be fixed. Applying J_+ to both sides we get

$$
(c_0+c_1)\,|\frac{1}{2},\frac{1}{2}\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle = 0 \tag{3.55}
$$

so $c_1 = -c_0$. Then on making the appropriate normalization we find

$$
|0,0\rangle = \frac{1}{\sqrt{2}}\big(\,|\frac{1}{2},-\frac{1}{2}\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle - |\frac{1}{2},\frac{1}{2}\rangle \otimes |\frac{1}{2},-\frac{1}{2}\rangle\,\big) \tag{3.56}
$$

Next, take the spin $1 \otimes$ spin $1/2$ composite system. As $j_1 = 1$, $j_2 = \frac{1}{2}$ there are two possible values for the composite spin, $j = \frac{1}{2}$ or $j = \frac{3}{2}$, and the tensor product space is 6-dimensional.

For the $j = \frac{3}{2}$ states, the state with greatest weight is

$$
|\frac{3}{2},\frac{3}{2}\rangle = |1,1\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle \tag{3.57}
$$

Applying J_- to both sides gives

$$|\frac{3}{2},\frac{1}{2}\rangle = \sqrt{\frac{2}{3}}\,|1,0\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle + \frac{1}{\sqrt{3}}\,|1,1\rangle \otimes |\frac{1}{2},-\frac{1}{2}\rangle \tag{3.58}$$

Further applications of J_- give

$$|\frac{3}{2},-\frac{1}{2}\rangle = \frac{1}{\sqrt{3}}\,|1,-1\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle + \sqrt{\frac{2}{3}}\,|1,0\rangle \otimes |\frac{1}{2},-\frac{1}{2}\rangle \tag{3.59}$$

and

$$|\frac{3}{2},-\frac{3}{2}\rangle = |1,-1\rangle \otimes |\frac{1}{2},-\frac{1}{2}\rangle \tag{3.60}$$

For the $j=\frac{1}{2}$ states, the state with maximal weight can be written as a linear combination

$$|\frac{1}{2},\frac{1}{2}\rangle = c_0\,|1,1\rangle \otimes |\frac{1}{2},-\frac{1}{2}\rangle + c_1\,|1,0\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle \tag{3.61}$$

for some constants c_0, c_1 to be determined. Then

$$0 = J_+\,|\frac{1}{2},\frac{1}{2}\rangle = (c_1 + \frac{1}{\sqrt{2}}c_0)\,|1,1\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle \tag{3.62}$$

so $c_1 = -\frac{1}{\sqrt{2}}c_0$. On making a unit normalization, we also fix c_0 and find

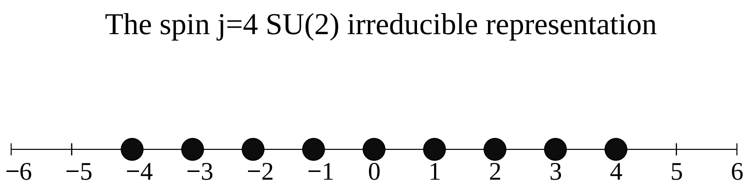
$$|\frac{1}{2},\frac{1}{2}\rangle = -\sqrt{\frac{2}{3}}\,|1,1\rangle \otimes |\frac{1}{2},-\frac{1}{2}\rangle + \frac{1}{\sqrt{3}}\,|1,0\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle \tag{3.63}$$

and applying J_- to both sides this gives

$$|\frac{1}{2},-\frac{1}{2}\rangle = -\frac{1}{\sqrt{3}}\,|1,0\rangle \otimes |\frac{1}{2},-\frac{1}{2}\rangle + \sqrt{\frac{2}{3}}\,|1,-1\rangle \otimes |\frac{1}{2},\frac{1}{2}\rangle \tag{3.64}$$

3.5 $SU(2)$ weight diagrams

One can plot the weights of an irreducible $\mathcal{L}(SU(2))$ representation on the real line; for example:



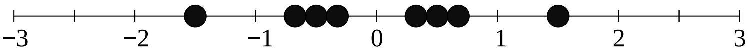
The weight diagrams of irreducible representations have the following properties

i) The diagram is reflection symmetric about the origin.

ii) The weights are all half integer; and are distributed with unit distance between each weight. The highest weight is j for $2j \in \mathbb{N}$, and the lowest weight is $-j$, and there are no "holes" in the weight diagram- $-j, -j+1, \ldots, j-1, j$ are all weights.

iii) Each weight has multiplicity 1.

One can also plot the weight diagram of a generic (not necessarily irreducible) representation. For example, the weight diagram of the tensor product $(j_1 = \frac{1}{2}) \otimes (j_2 = \frac{1}{2}) \otimes (j_3 = \frac{1}{2})$ eight dimensional tensor product representation is



This has a highest weight $j = \frac{3}{2}$ and a lowest weight $-\frac{3}{2}$ both with multiplicity 1, and weights $\pm\frac{1}{2}$ each with multiplicity 3. In general, a non-irreducible representation will have a highest weight, but it need not be of multiplicity 1. For a *generic* weight diagram

i) The diagram (together with weight multiplicities) is reflection symmetric about the origin.

ii) The weights are all half-integer.

iii) Each weight need not be of multiplicity 1. However, as one proceeds from a particular weight towards the origin (in unit steps from either the left or the right), the weight multiplicities do not decrease.

3.6 SU(2) in Particle Physics

3.6.1 Angular Momentum

The *orbital angular momentum* operators L_a acting on wavefunctions are given by

$$L_a = -i\epsilon_{abc} x^b \frac{\partial}{\partial x^c} \tag{3.65}$$

These operators satisfy

$$[L_a, L_b] = i\epsilon_{abc} L_c \tag{3.66}$$

and hence correspond to a (complexified) representation of $SU(2)$. Particles also carry a *spin angular momentum* $\mathbf{S}$, which commutes with the orbital angular momentum $[\mathbf{L}, \mathbf{S}] = 0$. The *total angular momentum* is defined by $\mathbf{J} = \mathbf{L} + \mathbf{S}$.

3.6.2 Isospin Symmetry

It is observed that the proton and neutron have similar mass, and also that the strong nuclear forces between nucleons are similar. Heisenberg introduced the concept of a $SU(2)$ isospin symmetry to systematize this. Particles are grouped into multiplets of isospin value I (previously called j) and labelled by the weights, which are the eigenvalues of I_3. Originally, this was formulated for nucleons, but later extended to describe all mesons and baryons.

Particles in the same isospin multiplet have the same baryon number, the same content of non-light quarks, the same spin and parity and almost the same mass. Isospin is a conserved quantum number in all known processes involving only strong interactions: it is related to the quark content by

$$I_3 = \frac{1}{2}(N(u) - N(\bar{u}) - (N(d) - N(\bar{d}))) \tag{3.67}$$

Isospin symmetry arises in the quark model because of the very similar properties of the u and d quarks.

Examples:

i) Nucleons have isospin $I = \frac{1}{2}$; the proton has $I_3 = \frac{1}{2}$, and the neutron has $I_3 = -\frac{1}{2}$:

$$\begin{aligned} n &= |\frac{1}{2}, -\frac{1}{2}\rangle \\ p &= |\frac{1}{2}, \frac{1}{2}\rangle \end{aligned} \tag{3.68}$$

ii) The pions have $I = 1$ with

$$\begin{aligned} \pi^- &= |1, -1\rangle \\ \pi^0 &= |1, 0\rangle \\ \pi^+ &= |1, 1\rangle \end{aligned} \tag{3.69}$$

iii) The strange baryons have $I = 0$ and $I = 1$

$$\begin{aligned} \Sigma^- &= |1, -1\rangle \\ \Sigma^0 &= |1, 0\rangle \\ \Sigma^+ &= |1, 1\rangle \end{aligned} \tag{3.70}$$

and

$$\Lambda^0 = |0, 0\rangle \tag{3.71}$$

iv) The strange mesons lie in two multiplets of $I = \frac{1}{2}$

$$\begin{aligned} K^0 &= |\frac{1}{2}, -\frac{1}{2}\rangle \\ K^+ &= |\frac{1}{2}, \frac{1}{2}\rangle \end{aligned} \tag{3.72}$$

and

$$\begin{aligned} K^- &= |\frac{1}{2}, -\frac{1}{2}\rangle \\ \bar{K}^0 &= |\frac{1}{2}, \frac{1}{2}\rangle \end{aligned} \tag{3.73}$$

$K^{\pm}$ are antiparticles with the same mass, but are in different isospin multiplets because of their differing quark content.

v) The light quarks have $I = \frac{1}{2}$

$$\begin{aligned} d &= |\frac{1}{2}, -\frac{1}{2}\rangle \\ u &= |\frac{1}{2}, \frac{1}{2}\rangle \end{aligned} \tag{3.74}$$

and all other quarks are isospin singlets $I = 0$.

3.6.3 Pauli's Generalized Exclusion Principle and the Deuteron

Consider first NN nucleon-nucleon bound states. From the previously obtained decomposition of the $(I_1 = 1/2) \otimes (I_2 = 1/2)$ tensor product we find the following isospin states

$$|1, 1\rangle = pp, \quad |1, 0\rangle = \frac{1}{\sqrt{2}}(np + pn), \quad |1, -1\rangle = nn \tag{3.75}$$

which are are symmetric under exchange of isospin degrees of freedom, and the remaining state is

$$|0, 0\rangle = \frac{1}{\sqrt{2}}(np - pn) \tag{3.76}$$

which is anti-symmetric under exchange of isospin degrees of freedom.

The deuteron d is a pn bound state, which has no pp or nn partners. There is therefore only one possibility; $d = |0, 0\rangle$.

In general, the total wavefunction for a NN state can be written as a product of space, spin, and isospin functions

$$\psi = \psi(\text{space})\psi(\text{spin})\psi(\text{isospin}) \tag{3.77}$$

The generalized Pauli exclusion principle requires allowed NN wave-functions be anti-symmetric under exchange of all degrees of freedom. As ψ(isospin) is anti-symmetric, ψ(space)ψ(spin) must be symmetric- in fact ψ(space) is symmetric (the nucleons in the d are in a $\ell = 0$ angular momentum state), and ψ(spin) as also symmetric, as d is a spin 1 particle.

3.6.4 Pion-Nucleon Scattering and Resonances

Isospin can be used to investigate scattering processes; consider for example pion-nucleon scattering processes. From the decomposition of the $(I_1 = 1) \otimes (I_2 = 1/2)$ tensor product we find

$$\begin{aligned} |\frac{3}{2}, \frac{3}{2}\rangle &= \pi^+ p \\ |\frac{3}{2}, \frac{1}{2}\rangle &= \sqrt{\frac{2}{3}}\pi^0 p + \frac{1}{\sqrt{3}}\pi^+ n \\ |\frac{3}{2}, -\frac{1}{2}\rangle &= \frac{1}{\sqrt{3}}\pi^- p + \sqrt{\frac{2}{3}}\pi^0 n \end{aligned}$$

$$|\frac{3}{2},-\frac{3}{2}\rangle = \pi^- n \tag{3.78}$$

and

$$\begin{aligned} |\frac{1}{2},\frac{1}{2}\rangle &= \frac{1}{\sqrt{3}}\pi^0 p - \sqrt{\frac{2}{3}}\pi^+ n \\ |\frac{1}{2},-\frac{1}{2}\rangle &= \sqrt{\frac{2}{3}}\pi^- p - \frac{1}{\sqrt{3}}\pi^0 n \end{aligned} \tag{3.79}$$

These equations can be inverted to give

$$\begin{aligned} \pi^+ p &= |\frac{3}{2},\frac{3}{2}\rangle \\ \pi^0 p &= \sqrt{\frac{2}{3}}\,|\frac{3}{2},\frac{1}{2}\rangle + \frac{1}{\sqrt{3}}\,|\frac{1}{2},\frac{1}{2}\rangle \\ \pi^+ n &= \frac{1}{\sqrt{3}}\,|\frac{3}{2},\frac{1}{2}\rangle - \sqrt{\frac{2}{3}}\,|\frac{1}{2},\frac{1}{2}\rangle \\ \pi^- p &= \frac{1}{\sqrt{3}}\,|\frac{3}{2},-\frac{1}{2}\rangle + \sqrt{\frac{2}{3}}\,|\frac{1}{2},-\frac{1}{2}\rangle \\ \pi^0 n &= \sqrt{\frac{2}{3}}\,|\frac{3}{2},-\frac{1}{2}\rangle - \frac{1}{\sqrt{3}}\,|\frac{1}{2},-\frac{1}{2}\rangle \\ \pi^- n &= |\frac{3}{2},-\frac{3}{2}\rangle \end{aligned} \tag{3.80}$$

Consider πN scattering. The scattering is described by a S-matrix, which, in processes dominated by strong interactions is taken to be isospin invariant: $[I_j, S] = 0$ and so by Schur's lemma

$$\langle I', m'|\, S\, |I, m\rangle = \phi(I)\delta_{II'}\delta_{mm'} \tag{3.81}$$

The cross sections of the scattering processes are given by

$$\sigma(\text{in} \to \text{out}) = K|\,\langle \text{in}|\, S\, |\text{out}\rangle\,|^2 \tag{3.82}$$

for constant K. Hence

$$\sigma(\pi^+ p \to \pi^+ p) = K|\phi(\frac{3}{2})|^2 \tag{3.83}$$

$$\sigma(\pi^0 n \to \pi^- p) = \frac{2}{9}K|\phi(\frac{3}{2}) - \phi(\frac{1}{2})|^2 \tag{3.84}$$

$$\sigma(\pi^- p \to \pi^- p) = \frac{1}{9}K|\phi(\frac{3}{2}) + 2\phi(\frac{1}{2})|^2 \tag{3.85}$$

For all three of these processes, a marked resonance is measured at approximately 1236 Mev. The ratio of the cross-sections of these resonances is

$$\sigma(\pi^+ p \to \pi^+ p) : \sigma(\pi^0 n \to \pi^- p) : \sigma(\pi^- p \to \pi^- p) = 1 : \frac{2}{9} : \frac{1}{9} \tag{3.86}$$

which is consistent with the supposition that the resonance corresponds to a particle of isospin $\frac{3}{2}$ (and so $|\phi(\frac{3}{2})| \gg |\phi(\frac{1}{2})|$). This particle is the Δ particle, which lies in an isospin $I = \frac{3}{2}$ multiplet with states Δ^-, Δ^0, Δ^+, Δ^{++} having weights $I_3 = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$ respectively.

3.7 The semi-simplicity of (complexified) $\mathcal{L}(SU(n+1))$

To proceed, define the $n+1 \times n+1$ square matrices E_{ij} by $(E_{i,j})_{pq} = \delta_{ip}\delta_{jq}$. Recall that $\mathcal{L}(SU(n+1))$ consists of the traceless antihermitian matrices, and is therefore spanned over $\mathbb{R}$ by $i(E_{i,i} - E_{n+1,n+1})$ for $i = 1, \dots, n$ and $E_{i,j} - E_{j,i}$, $i(E_{i,j} + E_{j,i})$ for $i < j$. Hence, on complexification, the Lie algebra is spanned by traceless diagonal matrices together with the $E_{i,j}$ for $i \neq j$.

Suppose that $h = \text{diag}\,(a_1, a_2, \dots, a_{n+1})$, $\sum a_i = 0$ is a traceless diagonal matrix. Then if $i \neq j$, observe that $[h, E_{i,j}] = (a_i - a_j)E_{i,j}$.

Proposition 25. *Complexified $\mathcal{L}(SU(n+1))$ is semi-simple.*
<u>Proof</u>

Note by direct computation that

$$[E_{i,j}, E_{r,s}] = \delta_{jr}E_{i,s} - \delta_{is}E_{r,j} \tag{3.87}$$

so

$$(\text{ad } E_{p,q} \text{ ad } E_{i,j})E_{r,s} = \delta_{jr}\delta_{iq}E_{p,s} - \delta_{jr}\delta_{ps}E_{i,q} - \delta_{is}\delta_{qr}E_{p,j} + \delta_{is}\delta_{pj}E_{r,q} \tag{3.88}$$

and hence the component of $(\text{ad } E_{p,q} \text{ ad } E_{i,j})E_{r,s}$ in the $E_{\ell,t}$ direction is

$$\delta_{jr}\delta_{iq}\delta_{p\ell}\delta_{st} - \delta_{jr}\delta_{ps}\delta_{i\ell}\delta_{qt} - \delta_{is}\delta_{qr}\delta_{p\ell}\delta_{jt} + \delta_{is}\delta_{pj}\delta_{r\ell}\delta_{qt} \tag{3.89}$$

There are various cases to consider

i) $\kappa(E_{p,q}, E_{i,j})$ for $p \neq q$, $i \neq j$. We must compute Tr $(\text{ad } E_{p,q} \text{ ad } E_{i,j})$. If $r \neq s$ then the component of $(\text{ad } E_{p,q} \text{ ad } E_{i,j})E_{r,s}$ in the $E_{r,s}$ direction is

$$\delta_{jr}\delta_{iq}\delta_{pr} - \delta_{jr}\delta_{ps}\delta_{ir}\delta_{qs} - \delta_{is}\delta_{qr}\delta_{pr}\delta_{js} + \delta_{is}\delta_{pj}\delta_{qs} = \delta_{jr}\delta_{iq}\delta_{pr} + \delta_{is}\delta_{pj}\delta_{qs} \tag{3.90}$$

(as $i \neq j$ and $p \neq q$). So the contribution to the trace from these terms is

$$\sum_{r=1,s=1,r\neq s}^{n+1} \delta_{jr}\delta_{iq}\delta_{pr} + \delta_{is}\delta_{pj}\delta_{qs} = 2n\delta_{iq}\delta_{pj} \tag{3.91}$$

We also compute the component of $(\text{ad } E_{p,q} \text{ ad } E_{i,j})E_{r,s}$ in the direction $E_{k,k} - E_{n+1,n+1}$ for $k = 1, \dots, n$. Observe that the component of $\delta_{iq}E_{p,s} - \delta_{ps}E_{iq}$ in this direction is $\delta_{iq}\delta_{ps}(\delta_{pk} - \delta_{ik})$ (if $i \neq q$ or $p \neq s$ then the diagonal components of $\delta_{iq}E_{p,s} - \delta_{ps}E_{iq}$ vanish). Hence the component of $(\text{ad } E_{p,q} \text{ ad } E_{i,j})E_{r,s}$ in the direction $E_{k,k} - E_{n+1,n+1}$ for $k = 1, \dots, n$ is

$$\delta_{jr}\delta_{iq}\delta_{ps}(\delta_{pk} - \delta_{ik}) + \delta_{is}\delta_{pj}\delta_{rq}(\delta_{rk} - \delta_{pk}) \tag{3.92}$$

It follows that the component of $(\text{ad } E_{p,q} \text{ ad } E_{i,j})(E_{k,k} - E_{n+1,n+1})$ along $(E_{k,k} - E_{n+1,n+1})$ is

$$\delta_{jk}\delta_{iq}\delta_{pk}(\delta_{pk} - \delta_{ik}) + \delta_{ik}\delta_{pj}\delta_{qk}(1 - \delta_{pk})$$

$$-\big(\delta_{j,n+1}\delta_{iq}\delta_{p,n+1}(\delta_{pk}-\delta_{ik})-\delta_{i,n+1}\delta_{pj}\delta_{q,n+1}\delta_{pk}\big) \tag{3.93}$$

which can be simplified using $i \neq j$ and $p \neq q$ to

$$\delta_{jk}\delta_{iq}\delta_{pk}+\delta_{ik}\delta_{pj}\delta_{qk}+\delta_{j,n+1}\delta_{iq}\delta_{p,n+1}\delta_{ik}+\delta_{i,n+1}\delta_{pj}\delta_{q,n+1}\delta_{pk} \tag{3.94}$$

On taking the sum from $k = 1, \ldots, n$ this gives a contribution to the trace of

$$2\delta_{iq}\delta_{pj} \tag{3.95}$$

and so

$$\kappa(E_{p,q}, E_{i,j}) = 2(n+1)\delta_{iq}\delta_{pj} \tag{3.96}$$

ii) $\kappa(E_{p,q}, h)$ for $h = \text{diag}\ (a_1, a_2, \ldots, a_{n+1})$ with $\sum_i a_i = 0$ and $p \neq q$. The only contribution to the trace Tr (ad $E_{p,q}$ ad h) comes from terms

$$(\text{ad } E_{p,q} \text{ ad } h)E_{i,j} = \text{ad } E_{p,q}(a_i-a_j)E_{i,j} = (a_i-a_j)(\delta_{qi}E_{p,j}-\delta_{p,j}E_{i,q}) \tag{3.97}$$

for $i \neq j$. The component of this matrix along the $E_{i,j}$ direction is

$$(a_i-a_j)(\delta_{qi}\delta_{pi}-\delta_{p,j}\delta_{q,j}) = 0 \tag{3.98}$$

as $p \neq q$. Hence

$$\kappa(E_{p,q}, h) = 0 \tag{3.99}$$

iii) $\kappa(h, g)$ where $h = \text{diag}\ (a_1, a_2, \ldots, a_{n+1})$ and $g = \text{diag}\ (b_1, b_2, \ldots, b_{n+1})$ and $\sum_i a_i = \sum_i b_i = 0$,

The only contribution to the trace Tr (ad h ad g) is from the terms

$$(\text{ad } h \text{ ad } g)E_{i,j} \tag{3.100}$$

for $i \neq j$. But

$$(\text{ad } h \text{ ad } g)E_{i,j} = (a_i-a_j)(b_i-b_j)E_{i,j} \tag{3.101}$$

so taking the sum over i and j $(i \neq j)$ we obtain

$$\kappa(h, g) = 2(n+1)\sum_i a_i b_i \tag{3.102}$$

Hence κ is negative definite over the span over $\mathbb{R}$ of $i(E_{r,r}-E_{n+1,n+1})$ for $r = 1, \ldots, n$; and this span is orthogonal to the span of the $E_{i,j}-E_{j,i}$ and $i(E_{i,j}+E_{j,i})$ $(i \neq j)$. Furthermore, κ is diagonal and negative definite over the span over $\mathbb{R}$ of the $E_{i,j}-E_{j,i}$ and $i(E_{i,j}+E_{j,i})$. κ is therefore non-degenerate. ■

We also have the immediate corollory:

Corollory 2. *The Lie algebra of $SU(n+1)$ (as a real Lie algebra) is compact.*

4. SU(3) and the Quark Model

The Lie algebra of $SU(3)$ consists of the traceless antihermitian 3×3 complex matrices. It is convenient to define the following matrices

$$
\begin{aligned}
h_1 &= \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} & h_2 &= \begin{pmatrix} \frac{1}{2\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{2\sqrt{3}} & 0 \\ 0 & 0 & -\frac{1}{\sqrt{3}} \end{pmatrix} \\
e^1_+ &= \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & e^1_- &= \begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
e^2_+ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 \end{pmatrix} & e^2_- &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \\
e^3_+ &= \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & e^3_- &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix}
\end{aligned} \tag{4.1}
$$

Then ih_1, ih_2 and $i(e^m_+ + e^m_-)$, $e^m_+ - e^m_-$ for $m = 1, 2, 3$ form a basis for the antihermitian traceless 3×3 matrices (over $\mathbb{R}$), and hence are a basis for $\mathcal{L}(SU(3))$. Suppose that d is the irreducible representation of $\mathcal{L}(SU(3))$ acting on a complex vector space V which is induced from an irreducible representation of $SU(3)$ acting on V.

It is convenient to set

$$
H_1 = d(h_1), \quad H_2 = d(h_2), \quad E^m_\pm = d(e^m_\pm) \text{ for } m = 1, 2, 3 \tag{4.2}
$$

Then we find the following commutators:

$$
\begin{aligned}
&[H_1, H_2] = 0 \\
&[H_1, E^1_\pm] = \pm E^1_\pm, \qquad [H_1, E^2_\pm] = \mp \frac{1}{2} E^2_\pm, \qquad [H_1, E^3_\pm] = \pm \frac{1}{2} E^3_\pm \\
&[H_2, E^1_\pm] = 0, \qquad [H_2, E^2_\pm] = \pm \frac{\sqrt{3}}{2} E^2_\pm, \qquad [H_2, E^3_\pm] = \pm \frac{\sqrt{3}}{2} E^3_\pm
\end{aligned} \tag{4.3}
$$

and

$$
\begin{aligned}
[E^1_+, E^1_-] &= H_1 \\
[E^2_+, E^2_-] &= \frac{\sqrt{3}}{2} H_2 - \frac{1}{2} H_1 \\
[E^3_+, E^3_-] &= \frac{\sqrt{3}}{2} H_2 + \frac{1}{2} H_1
\end{aligned} \tag{4.4}
$$

The remaining commutation relations are

$$
[E^1_+, E^2_+] = \frac{1}{\sqrt{2}} E^3_+, \qquad [E^1_-, E^2_-] = -\frac{1}{\sqrt{2}} E^3_-
$$

$$[E^1_+, E^3_-] = -\frac{1}{\sqrt{2}} E^2_-, \quad [E^1_-, E^3_+] = \frac{1}{\sqrt{2}} E^2_+$$
$$[E^2_+, E^3_-] = \frac{1}{\sqrt{2}} E^1_-, \qquad [E^2_-, E^3_+] = -\frac{1}{\sqrt{2}} E^1_+ \tag{4.5}$$

and

$$[E^1_+, E^2_-] = [E^1_-, E^2_+] = [E^1_+, E^3_+] = [E^1_-, E^3_-] = [E^2_+, E^3_+] = [E^2_-, E^3_-] = 0 \tag{4.6}$$

Note in particular that H_1, H_2 commute. The subalgebra of $\mathcal{L}(SU(3))$ spanned by ih_1 and ih_2 is called the Cartan subalgebra. It is the maximal commuting subalgebra of $\mathcal{L}(SU(3))$.

4.1 Raising and Lowering Operators: The Weight Diagram

The Lie algebra of $\mathcal{L}(SU(3))$ can be used to obtain three sets of $\mathcal{L}(SU(2))$ algebras. In particular, we find that

$$[H_1, E^1_\pm] = \pm E^1_\pm, \qquad [E^1_+, E^1_-] = H_1 \tag{4.7}$$

and

$$[\frac{\sqrt{3}}{2} H_2 - \frac{1}{2} H_1, E^2_\pm] = \pm E^2_\pm, \qquad [E^2_+, E^2_-] = \frac{\sqrt{3}}{2} H_2 - \frac{1}{2} H_1 \tag{4.8}$$

and

$$[\frac{\sqrt{3}}{2} H_2 + \frac{1}{2} H_1, E^3_\pm] = \pm E^3_\pm, \qquad [E^3_+, E^3_-] = \frac{\sqrt{3}}{2} H_2 + \frac{1}{2} H_1 \tag{4.9}$$

In particular, there are three pairs of raising and lowering operators $E^m_\pm$.

For simplicity, consider a representation d of $\mathcal{L}(SU(3))$ obtained from a unitary representation $\mathcal{D}$ of $SU(3)$ such that d is an anti-hermitian representation- so that H_1 and H_2 are hermitian, and hence diagonalizable with real eigenvalues. Hence, H_1 and $\frac{\sqrt{3}}{2}H_2 \pm \frac{1}{2}H_1$, can be simultaneously diagonalized, and the eigenvalues are real. (In fact the same can be shown without assuming unitarity!)

Suppose then that $|\phi\rangle$ is an eigenstate of H_1 with eigenvalue p and also an eigenstate of H_2 with eigenvalue q. It is convenient to order the eigenvalues as points in $\mathbb{R}^2$ with position vectors (p, q) where p is the eigenvalue of H_1 and q of H_2. (p, q) is then referred to as a weight.

From the commutation relations we have the following properties

i) Either $E^1_\pm |\phi\rangle = 0$ or $E^1_\pm |\phi\rangle$ is an eigenstate of H_1 and H_2 with eigenvalue $(p, q) \pm (1, 0)$

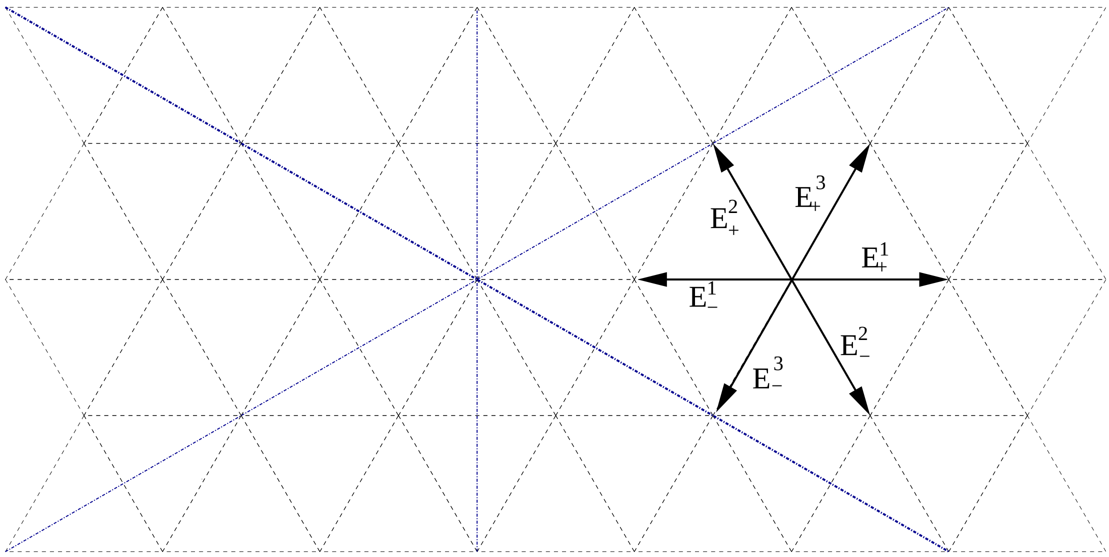
ii) Either $E^2_\pm |\phi\rangle = 0$ or $E^2_\pm |\phi\rangle$ is an eigenstate with eigenvalue $(p, q) \pm (-\frac{1}{2}, \frac{\sqrt{3}}{2})$

iii) Either $E^3_\pm |\phi\rangle = 0$ or $E^3_\pm |\phi\rangle$ is an eigenstate with eigenvalue $(p, q) \pm (\frac{1}{2}, \frac{\sqrt{3}}{2})$

Moreover, from the properties of $\mathcal{L}(SU(2))$ representations we know that

$$2p = m_1, \quad \sqrt{3}q - p = m_2, \quad \sqrt{3}q + p = m_3 \tag{4.10}$$

for m_1, m_2, $m_3 \in \mathbb{Z}$. It follows that $2\sqrt{3}q \in \mathbb{Z}$. It is particularly useful to plot the sets of eigenvalues (p, q) as points in the plane. The resulting plot is known as the weight diagram. As the representation is assumed to be irreducible, there can only be finitely many points on the weight diagram, though it is possible that a particular weight may correspond to more than one state. Moreover, as $2p \in \mathbb{Z}$, $2\sqrt{3}q \in \mathbb{Z}$, the weights are constrained to lie on the points of a lattice. From the effect of the raising and lowering operators on the eigenvalues, it is straightforward to see that this lattice is formed by the tessalation of the plane by equilateral triangles of side 1. This is illustrated in Figure 1, where the effect of the raising and lowering operators is given (in this diagram $(0, 0)$ is a weight, though this need not be the case generically).



The weight diagram has three axes of symmetry. To see this, recall that if m is a weight of a state in an irreducible representation of $\mathcal{L}(SU(2))$ then so is $-m$. In the context of the three $\mathcal{L}(SU(2))$ algebras contained in $\mathcal{L}(SU(3))$ this means that from the properties of the algebra in (4.7), if (p, q) is a weight then so is $(-p, q)$, i.e. the diagram is reflection symmetric about the line $\theta = \frac{\pi}{2}$ passing through the origin. Also, due to the symmetry of the $\mathcal{L}(SU(2))$ algebra in (4.8), the weight diagram is reflection symmetric about the line $\theta = \frac{\pi}{6}$ passing through the origin: so if (p, q) is a weight then so is $(\frac{1}{2}(p+\sqrt{3}q), \frac{1}{2}(\sqrt{3}p-q))$. And due to the symmetry of the $\mathcal{L}(SU(2))$ algebra in ((4.9) the weight diagram is reflection symmetric about the line $\theta = \frac{5\pi}{6}$ passing through the origin: so if (p, q) is a weight then so is $(\frac{1}{2}(p - \sqrt{3}q), \frac{1}{2}(-\sqrt{3}p - q))$.

Using this symmetry, it suffices to know the structure of the weight diagram in the sector of the plane $\frac{\pi}{6} \le \theta \le \frac{\pi}{2}$. The remainder is fixed by the reflection symmetry.

Motivated by the treatment of $SU(2)$ we make the definition:

Definition 42. *$|\psi\rangle$ is called a highest weight state if $|\psi\rangle$ is an eigenstate of both H_1 and H_2, and $E_+^m |\psi\rangle = 0$ for $m = 1, 2, 3$.*

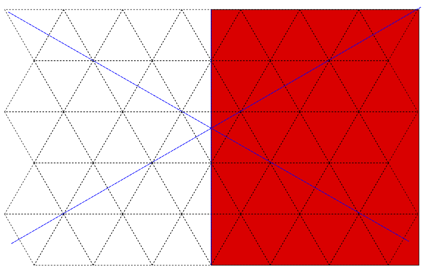
Note that there must be a highest weight state, for otherwise one could construct infinitely many eigenstates by repeated application of the raising operators E_+^m. Given a highest weight state, let V' be the vector space spanned by $|\psi\rangle$ and states obtained by acting with all possible products of lowering operators E_-^m on $|\psi\rangle$. As there are only finitely many points on the weight diagram, there can only be finitely many such terms. Then, by making use of the commutation relations, it is clear that V' is an invariant subspace of V. As the representation is irreducible on V, this implies that $V' = V$, i.e. V is spanned by $|\psi\rangle$ and a finite set of states obtained by acting with lowering operators on $|\psi\rangle$. Suppose that (p,q) is the weight of $|\psi\rangle$. Then V is spanned by a basis of eigenstates of H_1 and H_2 with weights confined to the sector given by $\pi \leq \theta \leq \frac{5\pi}{3}$ relative to (p,q)- all points on the weight diagram must therefore lie in this sector.

Lemma 6. *The highest weight state is unique.*

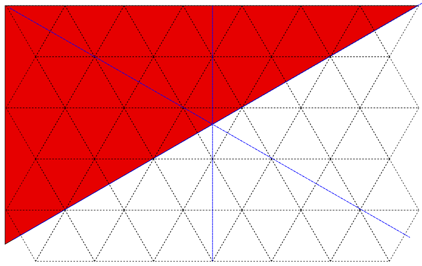
Proof Suppose that $|\psi\rangle$ and $|\psi'\rangle$ are two highest weight states with weights (p,q), (p',q') respectively. Then (p',q') must make an angle $\pi \leq \theta \leq \frac{5\pi}{3}$ relative to (p,q) and (p,q) must make an angle $\pi \leq \theta \leq \frac{5\pi}{3}$ relative to (p',q'). This implies that $p = p'$, $q = q'$.

Next suppose that $|\psi_1\rangle$ and $|\psi_2\rangle$ are two linearly independent highest weight states (both with weight (p,q)). Let V_1 and V_2 be the vector spaces spanned by the states obtained by acting with all possible products of lowering operators E_-^m on $|\psi_1\rangle$ and $|\psi_2\rangle$ respectively; one therefore obtains bases for V_1 and V_2 consisting of eigenstates of H_1 and H_2. By the reasoning given previously, as V_1 and V_2 are invariant subspaces of V and the representation is irreducible on V, it must be the case that $V_1 = V_2 = V$. In particular, we find that $|\psi_2\rangle \in V_1$. However, the only basis element of V_1 which has weight (p,q) is $|\psi_1\rangle$, hence we must have $|\psi_2\rangle = c\,|\psi_1\rangle$ for some constant c, in contradiction to the assumption that $|\psi_1\rangle$ and $|\psi_2\rangle$ are linearly independent. ■

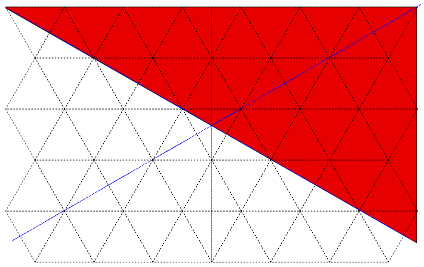
Having established the existence of a unique highest weight state $|\psi\rangle$, we can proceed to obtain the generic form for the weight diagram. Recall that the highest weight j of an irreducible representation of $\mathcal{L}(SU(2))$ is always non-negative. By acting on $|\psi\rangle$ with the lowering operators E_-^m, one obtains three irreducible representations of $\mathcal{L}(SU(2))$. Non-negativity of the highest weight corresponding to the $\mathcal{L}(SU(2))$ irreducible representation generated by E_-^1 implies that the highest weight must lie in the half-plane to the right of the line $\theta = \frac{\pi}{2}$, or on the line $\theta = \frac{\pi}{2}$:



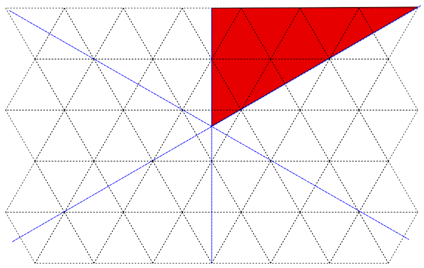
Non-negativity of the highest weight corresponding to the $\mathcal{L}(SU(2))$ irreducible representation generated by E_-^2 implies that the highest weight must lie in the half-plane above the line $\theta = \frac{\pi}{6}$, or on the line $\theta = \frac{\pi}{6}$:



Finally, non-negativity of the highest weight corresponding to the $\mathcal{L}(SU(2))$ irreducible representation generated by E_-^3 implies that the highest weight must lie in the half-plane above the line $\theta = \frac{5\pi}{6}$, or on the line $\theta = \frac{5\pi}{6}$:



As the highest weight must lie in all three of these regions, it must lie in the sector $\frac{\pi}{6} \leq \theta \leq \frac{\pi}{2}$ relative to $(0,0)$, or at the origin:



Lemma 7. *If the highest weight is* $(0,0)$, *then there is only one state in the representation, which is called the singlet.*

Proof

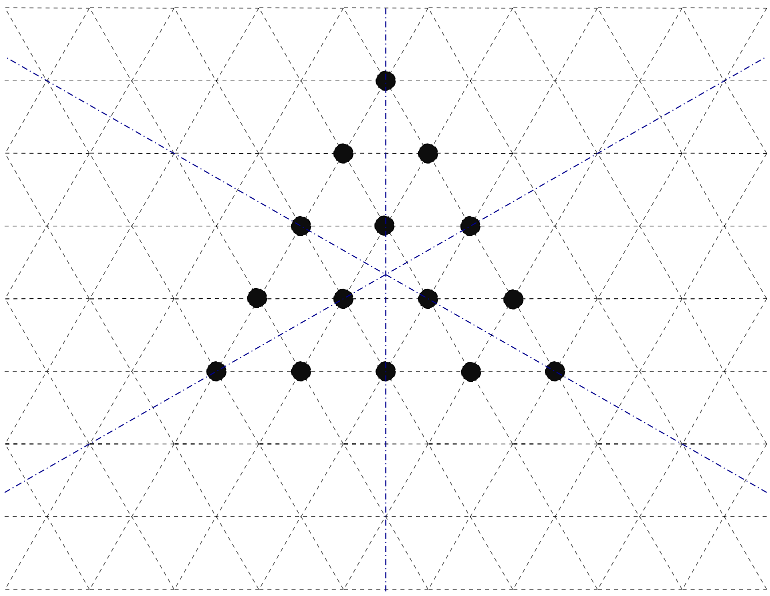
Let $|\psi\rangle$ be the highest weight state with weight $(0,0)$. Suppose that $E_-^m |\psi\rangle \neq 0$ for some m. Then by the reflection symmetry of the weight diagram, it follows that $E_+^m |\psi\rangle \neq 0$, in contradiction to the fact that $E_+^i |\psi\rangle = 0$ for $i = 1, 2, 3$, as $|\psi\rangle$ is the highest weight state. Hence $E_\pm^m |\psi\rangle = 0$ for $m = 1, 2, 3$. Also $H_1 |\psi\rangle = H_2 |\psi\rangle = 0$. It follows that the 1-dimensional subspace V' spanned by $|\psi\rangle$ is an invariant subspace of V, and therefore $V = V'$ as the representation is irreducible. ■

There are then three possible locations for the highest weight state $|\psi\rangle$.

4.1.1 Triangular Weight Diagrams (I)

Suppose that the highest weight lies on the line $\theta = \frac{\pi}{2}$. In this case, by applying powers of E_-^2 the states of the $\mathcal{L}(SU(2))$ representation given in (4.8) are generated. These form a line orthogonal to the axis of reflection $\theta = \frac{\pi}{6}$, about which they are symmetric, and there are no states outside this line, as these points cannot be reached by applying lowering operators. Then, by using the reflection symmetry, it follows that the outermost states from an equilateral triangle with horizontal base. Each lattice point inside the triangle corresponds to (at least) one state which has this weight, because each lattice point in the triangle lies at some possible weight within the $\mathcal{L}(SU(2))$ representation given in (4.7), and from the properties of $\mathcal{L}(SU(2))$ representations, we know that this has a state with this weight (i.e. as the $\mathcal{L}(SU(2))$ weight diagram has no "holes" in it, neither does the $\mathcal{L}(SU(3))$ weight diagram).

This case is illustrated by



Proposition 26. *Each weight in this triangle corresponds to a unique state.*

Proof

Note that all of the states on the right edge of the triangle correspond to unique states, because these weights correspond to states which can only be obtained by acting on $|\psi\rangle$ with powers of E^2_-. It therefore follows by the reflection symmetry that all of the states on the edges of the triangle have multiplicity one.

Now note the commutation relation

$$[E^1_-, E^2_-] = -\frac{1}{\sqrt{2}} E^3_- \tag{4.11}$$

This implies that products of lowering operators involving E^3_- can be rewritten as linear combinations of products of operators involving only E^1_- and E^2_- (in some order). In particular, we find

$$\begin{aligned} (E^1_-)(E^2_-)^n |\psi\rangle &= [E^1_-, E^2_-](E^2_-)^{n-1} |\psi\rangle + E^2_- E^1_-(E^2_-)^{n-1} |\psi\rangle \\ &= -\frac{1}{\sqrt{2}} E^3_-(E^2_-)^{n-1} |\psi\rangle + E^2_- E^1_-(E^2_-)^{n-1} |\psi\rangle \\ &\dots \end{aligned}$$

$$= -\frac{n}{\sqrt{2}} E^3_-(E^2_-)^{n-1} |\psi\rangle \tag{4.12}$$

by simple induction, where we have used the fact that $E^1_- |\psi\rangle = 0$ and $[E^2_-, E^3_-] = 0$.

A generic state of some fixed weight in the representation can be written as a linear combination of products of E^2_- and E^1_- lowering operators acting on $|\psi\rangle$ of the form

$$\Pi(E^1_-, E^2_-) |\psi\rangle \tag{4.13}$$

where $\Pi(E^1_-, E^2_-)$ contains m powers of E^2_- and ℓ powers of E^1_- where m, ℓ are uniquely determined by the weight of the state- only the order of the operators is unfixed.

Using (4.12), commute the E^1_- states in this product to the right as far as they will go. Then either one finds that the state vanishes (due to an E^1_- acting directly on $|\psi\rangle$), or one can eliminate all of the E^1_- terms and is left with a term proportional to

$$(E^2_-)^{m-\ell}(E^3_-)^{\ell} |\psi\rangle \tag{4.14}$$

where we have used the commutation relations $[E^2_-, E^3_-] = [E^1_-, E^3_-] = 0$.

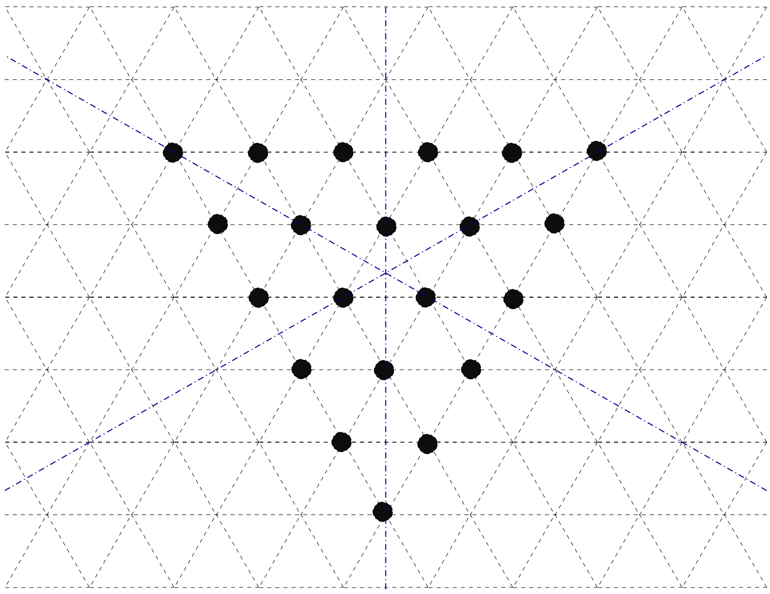
Hence, it follows that all weights in the diagram can have at most multiplicity 1. However, from the property of the $\mathcal{L}(SU(2))$ representations, as the weights in the outer layers have multiplicity 1, it follows that all weights in the interior have multiplicity at least 1.

Hence, all the weights must be multiplicity 1. ■

4.1.2 Triangular Weight Diagrams (II)

Suppose that the highest weight lies on the line $\theta = \frac{\pi}{6}$. In this case, by applying powers of E^1_- the states of the $\mathcal{L}(SU(2))$ representation given in (4.7) are generated. These form a horizontal line orthogonal to the axis of reflection $\theta = \frac{\pi}{2}$, about which they are symmetric, and there are no states outside this line, as these points cannot be reached by applying lowering operators. Then, by using the reflection symmetry, it follows that the outermost states from an inverted equilateral triangle with horizontal upper edge. Each lattice point inside the triangle corresponds to (at least) one state which has this weight, because each lattice point in the triangle lies at some possible weight within the $\mathcal{L}(SU(2))$ representation given in (4.7), and from the properties of $\mathcal{L}(SU(2))$ representations, we know that this has a state with this weight (i.e. as the $\mathcal{L}(SU(2))$ weight diagram has no "holes" in it, neither does the $\mathcal{L}(SU(3))$ weight diagram).

This case is illustrated by



Proposition 27. *Each weight in this triangle corresponds to a unique state.*

Proof

Note that all of the states on the horizontal top edge of the triangle correspond to unique states, because these weights correspond to states which can only be obtained by acting on $|\psi\rangle$ with powers of E^1_-. It therefore follows by the reflection symmetry that all of the states on the edges of the triangle have multiplicity one.

Now, using (4.11) it is straightforward to show that

$$E^2_-(E^1_-)^n\,|\psi\rangle = \frac{n}{\sqrt{2}}E^3_-(E^1_-)^{n-1}\,|\psi\rangle \tag{4.15}$$

for $n \geq 1$, where we have used $E^2_-\,|\psi\rangle = 0$.

Now consider a state of some fixed weight in the representation; this can be written as a linear combination of terms of the form

$$\Pi(E^1_-, E^2_-)\,|\psi\rangle \tag{4.16}$$

where $\Pi(E^1_-, E^2_-)$ contains m powers of E^1_- and ℓ powers of E^2_- in an appropriate order, where m and ℓ are determined uniquely by the weight of the state in question.

Using (4.15), commute the E^2_- states in this product to the right as far as they will go. Then either one finds that the state vanishes (due to an E^2_- acting directly on $|\psi\rangle$), or one can eliminate all of the E^1_- terms and is left with a term proportional to

$$(E^1_-)^{m-\ell}(E^3_-)^{\ell}\,|\psi\rangle \tag{4.17}$$

where we have used the commutation relations $[E^2_-, E^3_-] = [E^1_-, E^3_-] = 0$.

Hence, it follows that all weights in the diagram can have at most multiplicity 1. However, from the property of the $\mathcal{L}(SU(2))$ representations, as the weights in the outer layers have multiplicity 1, it follows that all weights in the interior have multiplicity at least 1.

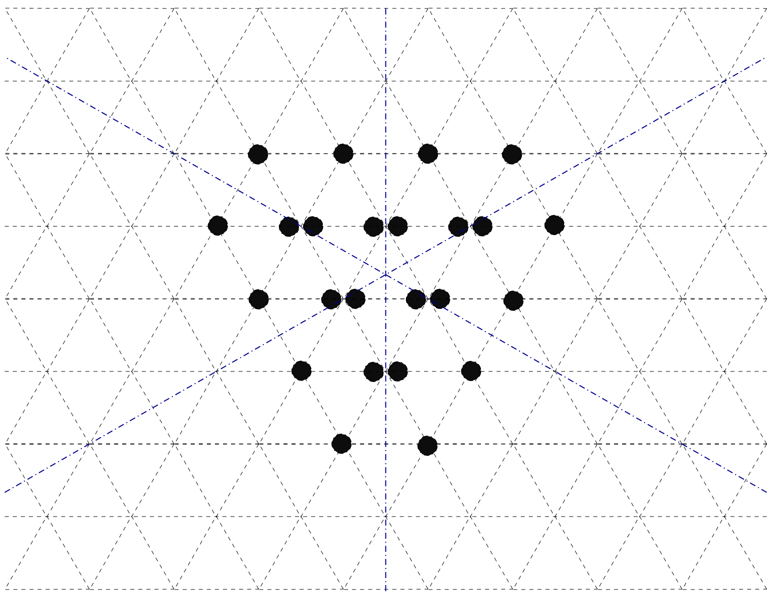
Hence, all the weights must be multiplicity 1. ■

4.1.3 Hexagonal Weight Diagrams

Suppose that the highest weight lies in the sector $\frac{\pi}{6} < \theta < \frac{\pi}{2}$. In this case, by applying powers of E_{-}^{1} the states of the $\mathcal{L}(SU(2))$ representation given in (4.7) are generated. These form a horizontal line extending to the left of the maximal weight which is orthogonal to the line $\theta = \frac{\pi}{2}$, about which they are symmetric, There are no states above, as these points cannot be reached by applying lowering operators. Also, by applying powers of E_{-}^{2} the states of the $\mathcal{L}(SU(2))$ representation given in (4.8) are generated. These form a line extending to the right of the maximal weight which is orthogonal to the axis of reflection $\theta = \frac{\pi}{6}$, about which they are symmetric, and there are no states to the right of this line, as these points cannot be reached by applying lowering operators.

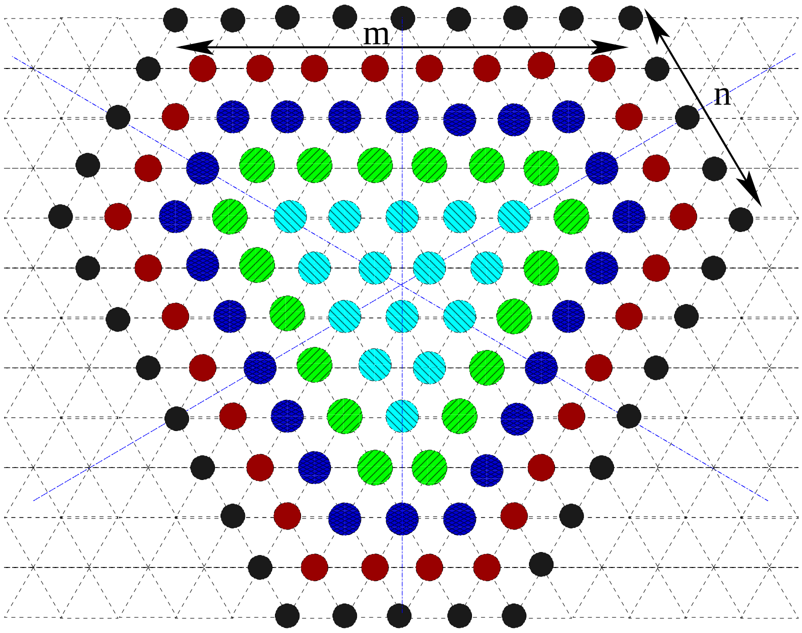
Then, by using the reflection symmetry, it follows that the outermost states form a hexagon. Each lattice point inside the hexagon corresponds to (at least) one state which has this weight, because each lattice point in the hexagon lies at some possible weight within the $\mathcal{L}(SU(2))$ representation given in (4.7), and from the properties of $\mathcal{L}(SU(2))$ representations, we know that this has a state with this weight (i.e. as the $\mathcal{L}(SU(2))$ weight diagram has no "holes" in it, neither does the $\mathcal{L}(SU(3))$ weight diagram).

This case is illustrated by



The multiplicities of the states for these weight diagrams are more complicated than for the triangular diagrams. In particular, the weights on the two edges of the hexagon leading off from the highest weight have multiplicity 1, because these states can only be constructed as $(E_{-}^{1})^{n} \left|\psi\right\rangle$ or $(E_{-}^{2})^{m} \left|\psi\right\rangle$. So by symmetry, all of the states on the outer layer of the hexagon have multiplicity 1. However, if one proceeds to the next layer, then the multiplicity of all the states increases by 1. This happens until the first triangular layer is reached, at which point all following layers have the same multiplicity as the first triangular layer.

Suppose that the top horizontal edge leading off the maximal weight is of length m, and that the other outer edge is of length n, with $m \geq n$. This situation is illustrated below



The highest weight is then at $(\frac{m}{2}, \frac{1}{2\sqrt{3}}(m+2n))$. The outer n layers are hexagonal, whereas the $n+1$-th layer is triangular, and all following layers are also triangular. As one goes inwards through the outer $n+1$ layers the multiplicity of the states in the layers increases from 1 in the first outer layer to $n+1$ in the $n+1$-th layer. Then all the states in the following triangular layers have multiplicity $n+1$ as well.

We will prove this in several steps.

Proposition 28. *States with weights on the k-th hexagonal layer for $k = 1, \ldots, n$ or the $k = n+1$-th layer (the first triangular layer) have multiplicity not exceeding k.*

Proof

In order to prove this, consider first a state on the upper horizontal edge of the k-th layer for $k \leq n+1$. The length of this edge is $m - k + 1$. A general state on this edge is obtained via

$$\Pi(E_-^2, E_-^1) \, |\psi\rangle \tag{4.18}$$

where $\Pi(E_-^2, E_-^1)$ contains (in some order) $k-1$ powers of E_-^2 and ℓ powers of E_-^1 for $\ell = k-1, \ldots, m$.

Now use the commutation relation (4.11) to commute the powers of E_-^2 to the right as far as they will go. Then the state can be written as a linear combination of the k vectors

$$|v_i\rangle = (E_-^3)^{i-1} (E_-^1)^{\ell-i+1} (E_-^2)^{k-i} \, |\psi\rangle \tag{4.19}$$

for $i = 1, \ldots, k$.

It follows that this state has multiplicity $\leq k$.

Next consider a state again on the k-th level, but now on the edge leading off to the right of the horizontal edge which we considered above; this edge is parallel to the outer edge of length n. Take $k \leq n+1$, so the edge has length $n-k+1$. A state on this edge is obtained via

$$\hat{\Pi}(E^1_-, E^2_-)\,|\psi\rangle \tag{4.20}$$

where $\hat{\Pi}(E^1_-, E^2_-)$ contains (in some order) $k-1$ powers of E^1_- and ℓ powers of E^2_- where $\ell = k-1, \ldots, n$. Now use the commutation relation (4.11) to commute the powers of E^1_- to the right as far as they will go. Then the state can be written as a linear combination of the k vectors

$$|w_i\rangle = (E^3_-)^{i-1}(E^2_-)^{\ell-i+1}(E^1_-)^{k-i}\,|\psi\rangle \tag{4.21}$$

for $i = 1, \ldots, k$.

So these states also have multiplicity $\leq k$. By using the reflection symmetry, it follows that the all the states on the k-th hexagonal layer have multiplicity k. ■

We also have the

Proposition 29. *The states with weights in the triangular layers have multiplicity not exceeding $n+1$.*

Proof

Consider a state on the k-th row of the weight diagram for $m+1 \geq k \geq n+1$ which lies inside the triangular layers. Such a state can also be written as

$$\Pi(E^2_-, E^1_-)\,|\psi\rangle \tag{4.22}$$

where $\Pi(E^2_-, E^1_-)$ contains (in some order) $k-1$ powers of E^2_- and ℓ powers of E^1_- for $\ell = k-1, \ldots, m$. and hence by the reasoning above, it can be rewritten as a linear combination of the k vectors $|v_i\rangle$ in (4.19), however for $i < k-n$, $|v_i\rangle = 0$ as $(E^2_-)^{k-i}\,|\psi\rangle = 0$. The only possible non-vanishing vectors are the $n+1$ vectors $|v_{k-n}\rangle, |v_{k-n+1}\rangle, \ldots, |v_k\rangle$. Hence these states have multiplicity $\leq n+1$. ■

Next note the lemma

Lemma 8. *Define $|w_{i,k}\rangle = (E^3_-)^{i-1}(E^1_-)^{k-i}(E^2_-)^{k-i}\,|\psi\rangle$ for $i = 1, \ldots, k$, $k = 1, \ldots, n+1$. Then the sets $S_k = \{|w_{1,k}\rangle, \ldots, |w_{k,k}\rangle\}$ are linearly independent for $k = 1, \ldots, n+1$.*

Proof

By using the commutation relations, it is straightforward to prove the identities

$$\begin{aligned}
E^3_+\,|w_{i,k}\rangle &= (i-1)\big(\frac{\sqrt{3}}{2}q + \frac{1}{2}p + \frac{i}{2} + 1 - k\big)\,|w_{i-1,k-1}\rangle \\
&\quad - \frac{1}{\sqrt{2}}(k-i)^2\big(\frac{\sqrt{3}}{2}q - \frac{1}{2}p + \frac{i}{2} + \frac{1}{2} - \frac{k}{2}\,|w_{i,k-1}\rangle \\
E^2_+\,|w_{i,k}\rangle &= E^1_-\big(\frac{1}{\sqrt{2}}(i-1)\,|w_{i-1,k-1}\rangle \\
&\quad + (k-i)\big(\frac{\sqrt{3}}{2}q - \frac{1}{2}p - \frac{1}{2}(k-i-1)\big)\,|w_{i,k-1}\rangle\,\big)
\end{aligned} \tag{4.23}$$

(with obvious simplifications in the cases when $i = 1$ or $i = k$)

Note that $S_1 = \{|\psi\rangle\}$ is linearly independent. Suppose that S_{k-1} is linearly independent ($k \geq 2$). Consider S_k. Suppose

$$\sum_{i=1}^{k} c_i \, |w_{i,k}\rangle = 0 \tag{4.24}$$

for some constants c_i. Applying E_+^3 to (4.24) and using the linear independence of S_{k-1} we find the relation

$$i(\frac{\sqrt{3}}{2} + \frac{1}{2}p + \frac{i}{2} + \frac{3}{2} - k)c_{i+1} - \frac{1}{\sqrt{2}}(k-i)^2(\frac{\sqrt{3}}{2}q - \frac{1}{2}p + \frac{i}{2} + \frac{1}{2} - \frac{1}{2}k)c_i = 0 \tag{4.25}$$

for $i = 1, \ldots, k-1$. Applying E_+^2 to (4.24) another recursion relation is obtained

$$\frac{1}{\sqrt{2}} i c_{i+1} + (k-i)(\frac{\sqrt{3}}{2}q - \frac{1}{2}p + \frac{i}{2} + \frac{1}{2} - \frac{1}{2}k)c_i = 0 \tag{4.26}$$

Combining these relations we find $c_{i+1} = 0$ for $i = 1, \ldots, k-1$. If $\frac{\sqrt{3}}{2}q - \frac{1}{2}p + \frac{i}{2} + \frac{1}{2} - \frac{1}{2}k \neq 0$ when $i = 1$ then one also has $c_1 = 0$. This holds if $k \leq n+1$, however if $k = n+2$ then c_1 is not fixed by these equations. The induction stops at this point. ■

These results are sufficient to fix the multiplicity of all the states. This is because the vectors in S_k for $1 \leq k \leq k+1$ correspond to states with weight $(p, q) - (k-1)(\frac{1}{2}, \frac{\sqrt{3}}{2})$ which are at the top right hand corner of the k-th hexagonal (or outermost triangular for $k = n+1$) layer. We have shown therefore that these weights have multiplicity both less than or equal to, and greater than or equal to k. Hence these weights have multiplicity k. Next consider the states on the level k edges which are obtained by acting with the $\mathcal{L}(SU(2))$ lowering operators E_-^1 and E_-^2 on the "corner weight" states. Observe the following:

Lemma 9. *Let d be a representation of $\mathcal{L}(SU(2))$ on V be such that a particular $\mathcal{L}(SU(2))$ weight $m > 0$ has multiplicity p. Then all weights m' such that $|m'| \leq m$ have multiplicity $\geq p$*

whose proof is left as an exercise.

By this lemma, all the states on the k-th layer obtained in this fashion have multiplicity k also. Then the reflection symmetry implies that all states on the k-th layer have multiplicity k.

In particular, the states on the outer triangular layer have multiplicity $n+1$. We have shown that the states on the triangular layers must have multiplicity not greater than $n+1$, but by the lemma above together with the reflection symmetry, they must also have multiplicity $\geq n+1$. Hence the triangular layer weights have multiplicity $n+1$, and the proof is complete.

This was rather long-winded. There exist general formulae constraining multiplicities of weights in more general Lie group representations, but we will not discuss these here.

4.1.4 Dimension of Irreducible Representations

Using the multiplicity properties of the weight diagram, it is possible to compute the dimension of the representation. We consider first the hexagonal weight diagram for $m \geq n$.

Then there are $1 + \cdots + (m-n) + (m-n+1) = \frac{1}{2}(m-n+1)(m-n+2)$ weights in the interior triangle. Each of these weights has multiplicity $(n+1)$ which gives $\frac{1}{2}(n+1)(m-n+1)(m-n+2)$ linearly independent states corresponding to weights in the triangle. Consider next the k-th hexagonal layer for $k = 1, \ldots, n$. This has $3((m+1-(k-1)) + (n+1-(k-1)) - 2) = 3(m+n+2-2k)$ weights in it, and each weight has multiplicity k, which gives $3k(m+n+2-2k)$ linearly independent states in the k-th hexagonal layer.

The total number of linearly independent states is then given by

$$\frac{1}{2}(n+1)(m-n+1)(m-n+2) + \sum_{k=1}^{n} 3k(m+n+2-2k) = \frac{1}{2}(m+1)(n+1)(m+n+2) \quad (4.27)$$

This formula also applies in the case for $m \leq n$ and also for the triangular weight diagrams by taking $m = 0$ or $n = 0$. The lowest dimensional representations are therefore 1,3,6,8,10...

4.1.5 The Complex Conjugate Representation

Definition 43. *Let d be a representation of a Lie algebra $\mathcal{L}(G)$ acting on V. If $v \in \mathcal{L}(G)$, then viewing $d(v)$ as a matrix acting on V, the complex representation $\bar{d}$ is defined by*

$$\bar{d}(v)u = (d(v))^* u \quad (4.28)$$

for $u \in V$, where $*$ denotes matrix complex conjugation.

Note that as $d(v)$ is linear in v over $\mathbb{R}$, it follows that $(d(v))^*$ is also linear in v over $\mathbb{R}$. Also, as

$$d([v,w]) = d(v)d(w) - d(w)d(v) \quad (4.29)$$

for $v, w \in \mathcal{L}(G)$, so taking the complex conjugate of both sides we find

$$\bar{d}([v,w]) = \bar{d}(v)\bar{d}(w) - \bar{d}(w)\bar{d}(v) \quad (4.30)$$

i.e. $\bar{d}$ is indeed a Lie algebra representation.

Suppose that T_a are the generators of $\mathcal{L}(G)$ with structure constants $c_{ab}{}^c$. Then as d is a representation,

$$[d(T_a), d(T_b)] = c_{ab}{}^c d(T_c) \quad (4.31)$$

Taking the complex conjugate, and recalling that $c_{ab}{}^c$ are real, we find

$$[\bar{d}(T_a), \bar{d}(T_b)] = c_{ab}{}^c \bar{d}(T_c) \quad (4.32)$$

i.e. the $d(T_a)$ and $\bar{d}(T_a)$ satisfy the same commutation relations.

In the context of representations of $\mathcal{L}(SU(3))$, the conjugate operators to iH_1, iH_2, $i(E_+^m + E_-^m)$ and $E_+^m - E_-^m$ are denoted by $i\bar{H}_1$, $i\bar{H}_2$, $i(\bar{E}_-^m + \bar{E}_+^m)$, and $\bar{E}_+^m - \bar{E}_-^m$ respectively and are given by

$$i\bar{H}_1 = (iH_1)^*$$

$$\begin{aligned} i\bar{H}_2 &= (iH_2)^* \\ i(\bar{E}^m_- + \bar{E}^m_+) &= (i(E^m_+ + E^m_-))^* \\ \bar{E}^m_+ - \bar{E}^m_- &= (E^m_+ - E^m_-)^* \end{aligned} \tag{4.33}$$

which implies

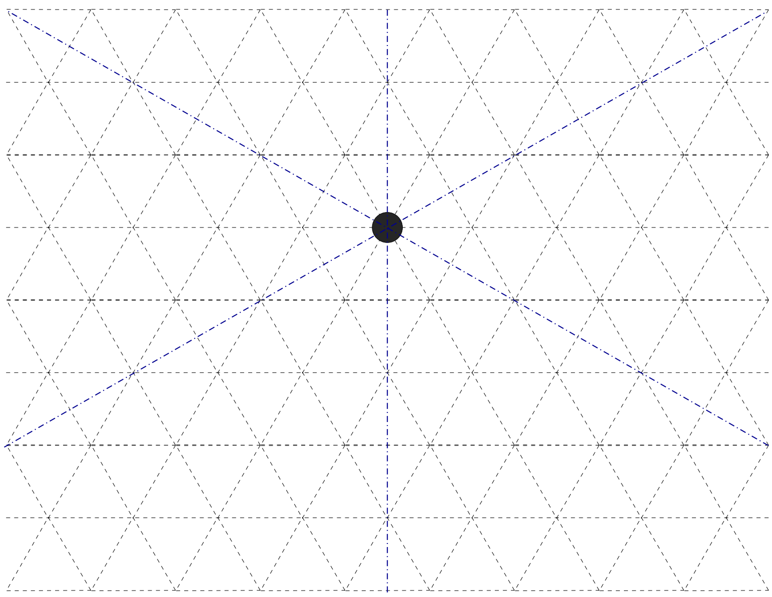
$$\bar{H}_1 = -(H_1)^*, \quad \bar{H}_2 = -(H_2)^*, \quad \bar{E}^m_\pm = -(E^m_\mp)^* \tag{4.34}$$

Then $\bar{H}_1$, $\bar{H}_2$ and $\bar{E}^m_\pm$ satisfy the same commutation relations as the unbarred operators, and also behave in the same way under the hermitian conjugate. One can therefore plot the weight diagram associated with the conjugate representation $\bar{d}$, the weights being the (real) eigenvalues of $\bar{H}_1$ and $\bar{H}_2$. But as $\bar{H}_1 = -(H_1)^*$ and $\bar{H}_2 = -(H_2)^*$ it follows that if (p,q) is a weight of the representation d, then $(-p,-q)$ is a weight of the representation $\bar{d}$. So the weight diagram of $\bar{d}$ is obtained from that of d by inverting all the points $(p,q) \to -(p,q)$. Note that this means that the equilateral triangular weight diagrams ▲ and ▼ of equal length sides are conjugate to each other.

4.2 Some Low-Dimensional Irreducible Representations of $\mathcal{L}(SU(3))$

4.2.1 The Singlet

The simplest representation has only one state, which is the highest weight state with weight $(0,0)$. This representation is denoted **1**.

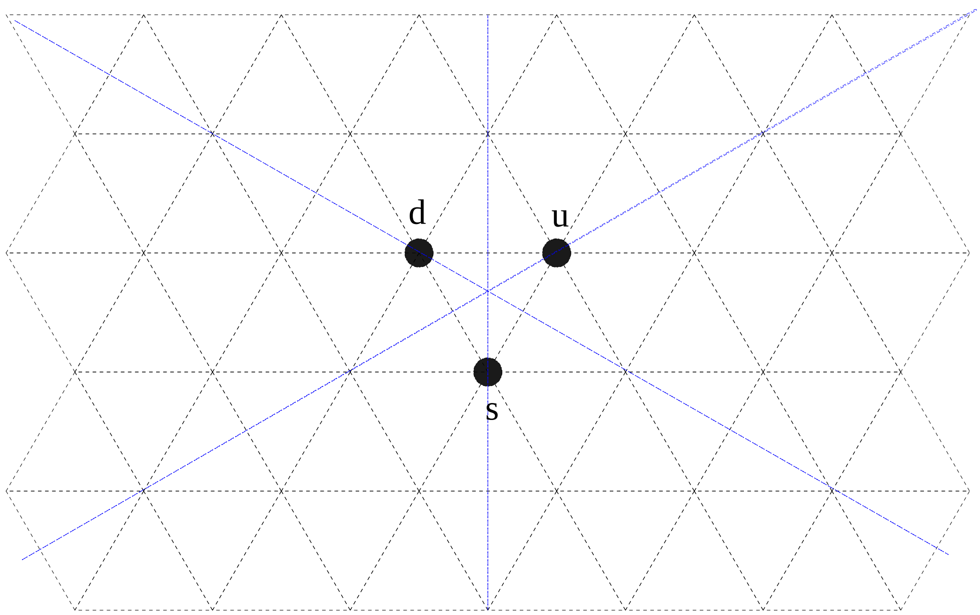


4.2.2 3-dimensional Representations

Take the fundamental representation. Then as h_1 and h_2 are already diagonalized, it is straightforward to compute the eigenstates and weights.

State	*Weight*
$\begin{pmatrix}1\\0\\0\end{pmatrix}$	$(\frac{1}{2}, \frac{1}{2\sqrt{3}})$
$\begin{pmatrix}0\\1\\0\end{pmatrix}$	$(-\frac{1}{2}, \frac{1}{2\sqrt{3}})$
$\begin{pmatrix}0\\0\\1\end{pmatrix}$	$(0, -\frac{1}{\sqrt{3}})$

The state of highest weight is $\begin{pmatrix}1\\0\\0\end{pmatrix}$ which has weight $(\frac{1}{2}, \frac{1}{2\sqrt{3}})$. The weight diagram is



This representation is denoted **3**. It will be convenient to define the following states in the **3** representation.

$$u = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad d = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad s = \begin{pmatrix}0\\0\\1\end{pmatrix} \tag{4.35}$$

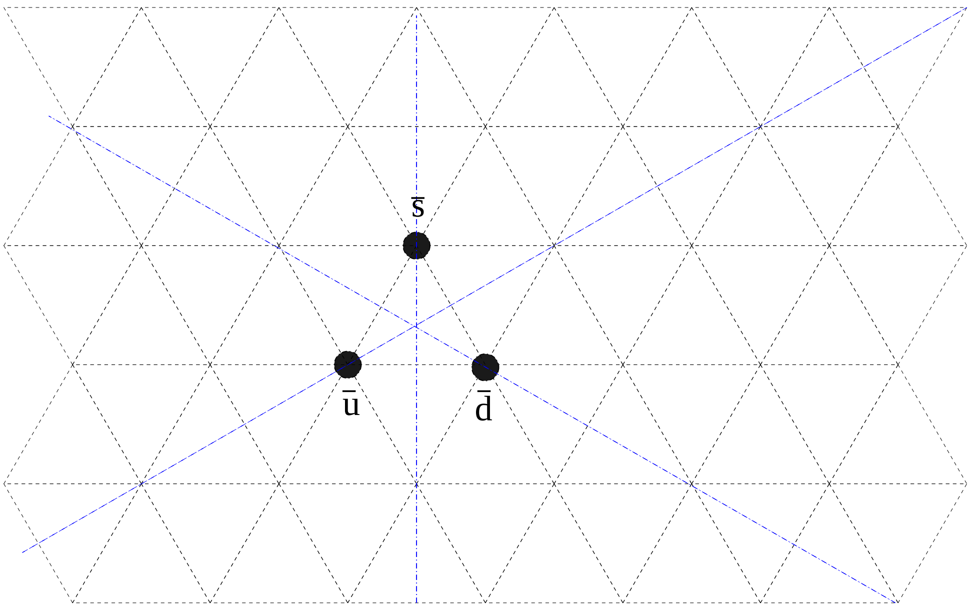
so that u has weight $(\frac{1}{2}, \frac{1}{2\sqrt{3}})$, d has weight $(-\frac{1}{2}, \frac{1}{2\sqrt{3}})$ and s has weight $(0, -\frac{1}{\sqrt{3}})$.

The lowering operators have the following effect: $d = \sqrt{2}e^1_- u$, $s = \sqrt{2}e^3_- u$ and $s = \sqrt{2}e^2_- d$.

The complex conjugate of this representation is called $\bar{\mathbf{3}}$ and the weights are obtained by multiplying the weights of the **3** representation by -1.

State	*Weight*
$\begin{pmatrix}1\\0\\0\end{pmatrix}$	$(-\frac{1}{2}, -\frac{1}{2\sqrt{3}})$
$\begin{pmatrix}0\\1\\0\end{pmatrix}$	$(\frac{1}{2}, -\frac{1}{2\sqrt{3}})$
$\begin{pmatrix}0\\0\\1\end{pmatrix}$	$(0, \frac{1}{\sqrt{3}})$

The state of highest weight is $\begin{pmatrix}0\\0\\1\end{pmatrix}$ which has weight $(0, \frac{1}{\sqrt{3}})$. The weight diagram is



It will be convenient to define the following states in the $\bar{\mathbf{3}}$ representation.

$$\bar{u} = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \bar{d} = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad \bar{s} = \begin{pmatrix}0\\0\\1\end{pmatrix} \tag{4.36}$$

so that $\bar{u}$ has weight $(-\frac{1}{2}, -\frac{1}{2\sqrt{3}})$, $\bar{d}$ has weight $(\frac{1}{2}, -\frac{1}{2\sqrt{3}})$ and $\bar{s}$ has weight $(0, \frac{1}{\sqrt{3}})$.

The lowering operators have the following effect: $\bar{u} = -\sqrt{2}\bar{e}^3_-\bar{s}$, $\bar{d} = -\sqrt{2}\bar{e}^2_-\bar{s}$ and $\bar{u} = -\sqrt{2}\bar{e}^1_-\bar{d}$; where $\bar{e}^m_\pm = -(e^m_\mp)^*$.

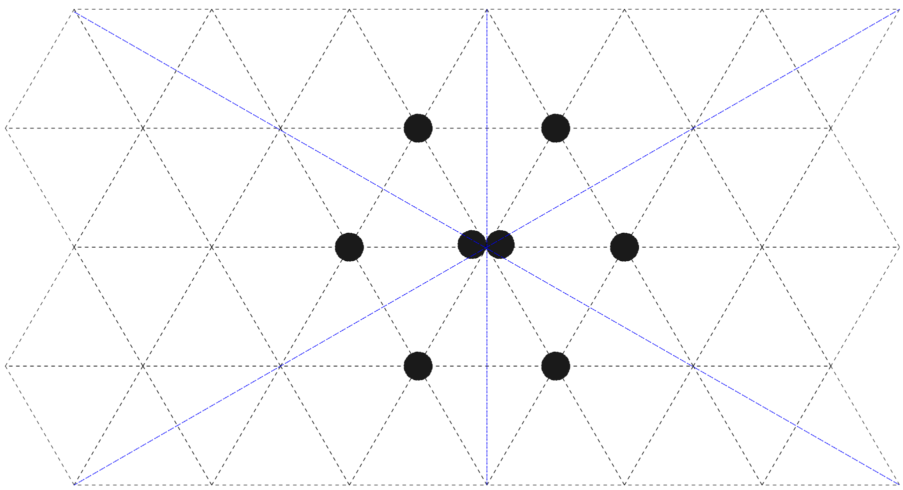
Exercise: Verify that all other lowering operators $\bar{e}^m_-$ (except those given above) annihilate $\bar{u}$, $\bar{d}$, $\bar{s}$. Also compute the effect of the raising operators $\bar{e}^m_+$.

4.2.3 Eight-Dimensional Representations

Consider the adjoint representation defined on the *complexified* Lie algebra $\mathcal{L}(SU(3))$, i.e. $\mathrm{ad}(v)w = [v, w]$. Then the weights of the states can be computed by evaluating the commutators with h_1 and h_2:

State v	$[h_1, v]$	$[h_2, v]$	*Weight*
h_1	0	0	$(0, 0)$
h_2	0	0	$(0, 0)$
e^1_+	e^1_+	0	$(1, 0)$
e^1_-	$-e^1_-$	0	$(-1, 0)$
e^2_+	$-\frac{1}{2}e^2_+$	$\frac{\sqrt{3}}{2}e^2_+$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$
e^2_-	$\frac{1}{2}e^2_-$	$-\frac{\sqrt{3}}{2}e^2_-$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$
e^3_+	$\frac{1}{2}e^3_+$	$\frac{\sqrt{3}}{2}e^3_+$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
e^3_-	$-\frac{1}{2}e^3_-$	$-\frac{\sqrt{3}}{2}e^3_-$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$

The highest weight state is e^3_+ with weight $(\frac{1}{2}, \frac{\sqrt{3}}{2})$. All weights have multiplicity 1 except for $(0, 0)$ which has multiplicity 2. The weight diagram is a regular hexagon:



4.3 Tensor Product Representations

Suppose that d_1, d_2 are irreducible representations of $\mathcal{L}(SU(3))$ acting on V_1, V_2 respectively. Then let $V = V_1 \bigotimes V_2$ and $d = d_1 \otimes 1 + 1 \otimes d_2$ be the tensor product representation of $\mathcal{L}(SU(3))$ on V. In general d is not irreducible on V, so we want to decompose V into a direct sum of invariant subspaces on which the restriction of d is irreducible.

To do this, recall that one can choose a basis of V_1 which consists entirely of eigenstates of both $d_1(h_1)$ and $d_1(h_2)$. Similarly, one can also choose a basis of V_2 which consists entirely of eigenstates of both $d_2(h_1)$ and $d_2(h_2)$. Then the tensor product of the basis eigenstates produces a basis of $V_1 \bigotimes V_2$ which consists of eigenstates of $d(h_1)$ and $d(h_2)$.

Explicitly, suppose that $|\phi_i\rangle \in V_i$ is an eigenstate of $d_i(h_1)$ and $d_i(h_2)$ with weight (p_i, q_i) (i.e. $d_i(h_1)|\phi_i\rangle = p_i|\phi_i\rangle$ and $d_i(h_2)|\phi_i\rangle = q_i|\phi_i\rangle$) for $i = 1, 2$. Define $|\phi\rangle =$

$|\phi_1\rangle \otimes |\phi_2\rangle$. Then

$$\begin{aligned} d(h_1)\,|\phi\rangle &= (d_1(h_1)\,|\phi_1\rangle) \otimes |\phi_2\rangle + |\phi_1\rangle \otimes (d_2(h_1)\,|\phi_2\rangle) \\ &= (p_1\,|\phi_1\rangle) \otimes |\phi_2\rangle + |\phi_1\rangle \otimes (p_2\,|\phi_2\rangle) \\ &= (p_1 + p_2)\,|\phi\rangle \end{aligned} \tag{4.37}$$

and similarly

$$d(h_2)\,|\phi\rangle = (q_1 + q_2)\,|\phi\rangle \tag{4.38}$$

So the weight of $|\phi\rangle$ is $(p_1 + p_2, q_1 + q_2)$; the weights add in the tensor product representation.

Using this, one can plot a weight diagram consisting of the weights of all the eigenstates in the tensor product basis of V, the points in the weight diagram are obtained by adding the pairs of weights from the weight diagrams of d_1 and d_2 respectively, keeping careful track of the multiplicities (as the same point in the tensor product weight diagram may be obtained from adding weights from different states in $V_1 \bigotimes V_2$.)

Once the tensor product weight diagram is constructed, pick a highest weight, which corresponds to a state which is annihilated by the tensor product operators E_+^m for $m = 1, 2, 3$. (Note that as the representation is finite-dimensional such a state is guaranteed to exist, though as the representation is no longer irreducible, it need not be unique). If there are multiple highest weight states corresponding to the same highest weight, one can without loss of generality take them to be mutually orthogonal. Picking one of these, generate further states by acting on a highest weight state with all possible combinations of lowering operators. The span of these (finite number) of states produces an invariant subspace W_1 of V on which d is irreducible. Remove these weights from the tensor product weight diagram. If the multiplicity of one of the weights in the original tensor product diagram is k, and the multiplicity of the same weight in the W_1 weight diagram is k' then on removing the W_1 weights, the multiplicity of that weight must be reduced from k to $k - k'$.

Repeat this process until there are no more weights left. This produces a decomposition $V = W_1 \bigoplus ... \bigoplus W_k$ of V into invariant subspaces W_j on which d is irreducible.

Note that one could also perform this process on triple (and higher order) tensor products e.g. $V_1 \bigotimes V_2 \bigotimes V_3$. In this case, one would construct the tensor product weight diagram by adding triplets of weights from the weight diagrams of d_1 on V_1, d_2 on V_2 and d_3 on V_3 respectively.

This process can be done entirely using the weight diagrams, because we have shown that for irreducible representations, the location of the highest weight fixes uniquely the shape of the weight diagram and the multiplicities of its states.

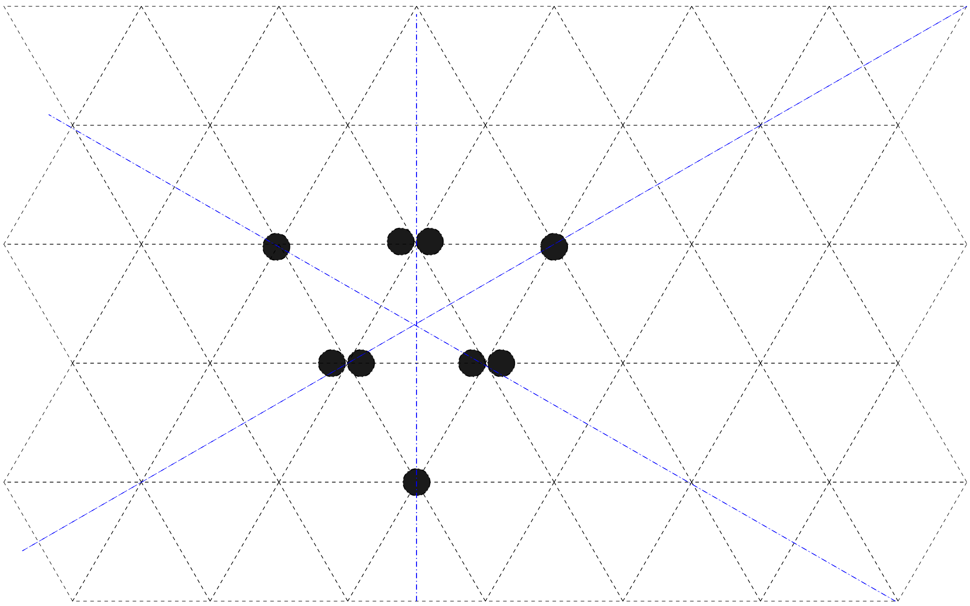
We will see how this works for some simple examples:

4.3.1 $\mathbf{3} \otimes \mathbf{3}$ decomposition.

Consider the $\mathbf{3} \otimes \mathbf{3}$ tensor product. Adding the weights together one obtains the following table of quark content and associated weights

Quark content and weights for $\mathbf{3} \otimes \mathbf{3}$	
Quark Content	*Weight*
$u \otimes u$	$(1, \frac{1}{\sqrt{3}})$
$d \otimes d$	$(-1, \frac{1}{\sqrt{3}})$
$s \otimes s$	$(0, -\frac{2}{\sqrt{3}})$
$u \otimes d$, $d \otimes u$	$(0, \frac{1}{\sqrt{3}})$
$u \otimes s$, $s \otimes u$	$(\frac{1}{2}, -\frac{1}{2\sqrt{3}})$
$d \otimes s$, $s \otimes d$	$(-\frac{1}{2}, -\frac{1}{2\sqrt{3}})$

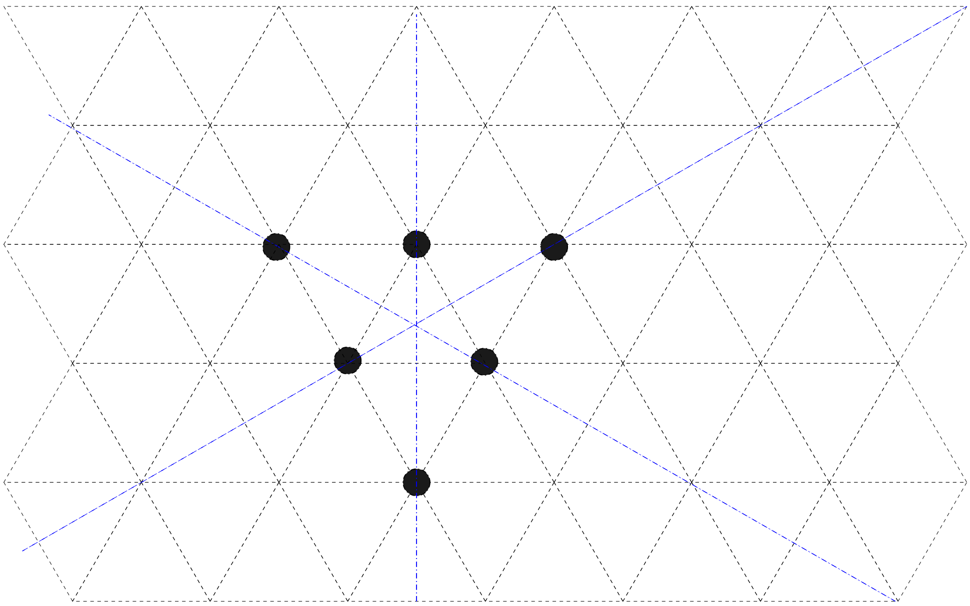
Plotting the corresponding weight diagram gives



The raising and lowering operators are $E^m_\pm = e^m_\pm \otimes 1 + 1 \otimes e^m_\pm$. The highest weight state is $u \otimes u$ with weight $(1, \frac{1}{\sqrt{3}})$. Applying lowering operators to $u \otimes u$ it is clear that a six-dimensional irreducible representation is obtained. The (unit-normalized) states and weights are given by

States and weights for the $\mathbf{6}$ in $\mathbf{3} \otimes \mathbf{3}$	
State	*Weight*
$u \otimes u$	$(1, \frac{1}{\sqrt{3}})$
$d \otimes d$	$(-1, \frac{1}{\sqrt{3}})$
$s \otimes s$	$(0, -\frac{2}{\sqrt{3}})$
$\frac{1}{\sqrt{2}}(d \otimes u + u \otimes d)$	$(0, \frac{1}{\sqrt{3}})$
$\frac{1}{\sqrt{2}}(u \otimes s + s \otimes u)$	$(\frac{1}{2}, -\frac{1}{2\sqrt{3}})$
$\frac{1}{\sqrt{2}}(d \otimes s + s \otimes d)$	$(-\frac{1}{2}, -\frac{1}{2\sqrt{3}})$

which has the following weight diagram



This representation is called **6**. Removing the (non-vanishing) span of these states from the tensor product space, one is left with a 3-dimensional vector space. The new highest weight is at $(0, \frac{1}{\sqrt{3}})$ with corresponding state $\frac{1}{\sqrt{2}}(d \otimes u - u \otimes d)$ (this is the unique linear combination- up to overall scale- of $d \otimes u$ and $u \otimes d$ which is annihilated by all the raising operators). This generates a $\bar{\mathbf{3}}$. The states and their weights are

States and weights for the $\bar{\mathbf{3}}$ in $\mathbf{3} \otimes \mathbf{3}$	
State	*Weight*
$\frac{1}{\sqrt{2}}(d \otimes u - u \otimes d)$	$(0, \frac{1}{\sqrt{3}})$
$\frac{1}{\sqrt{2}}(d \otimes s - s \otimes d)$	$(-\frac{1}{2}, -\frac{1}{2\sqrt{3}})$
$\frac{1}{\sqrt{2}}(s \otimes u - u \otimes s)$	$(\frac{1}{2}, -\frac{1}{2\sqrt{3}})$

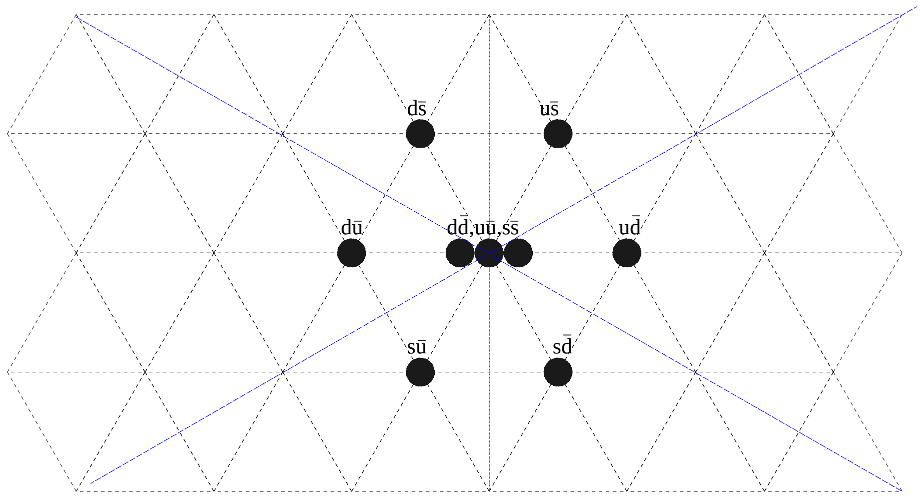
Hence $\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \bar{\mathbf{3}}$. The states in the **6** are symmetric, whereas those in the $\bar{\mathbf{3}}$ are antisymmetric.

4.3.2 $\mathbf{3} \otimes \bar{\mathbf{3}}$ decomposition

For this tensor product the quark content/weight table is as follows:

Quark content and weights for $\mathbf{3} \otimes \bar{\mathbf{3}}$	
Quark Content	*Weight*
$u \otimes \bar{s}$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
$u \otimes \bar{d}$	$(1, 0)$
$d \otimes \bar{s}$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$
$u \otimes \bar{u}$, $d \otimes \bar{d}$, $s \otimes \bar{s}$	$(0, 0)$
$d \otimes \bar{u}$	$(-1, 0)$
$s \otimes \bar{u}$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$s \otimes \bar{d}$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$

with weight diagram



The raising and lowering operators are $E^m_\pm = e^m_\pm \otimes 1 + 1 \otimes \bar{e}^m_\pm$ All weights have multiplicity 1, except for $(0, 0)$ which has multiplicity 3. The highest weight state is $u \otimes \bar{s}$ with weight $(\frac{1}{2}, \frac{\sqrt{3}}{2})$. Acting on this state with all possible lowering operators one obtains an $\mathbf{8}$ with the following states and weights

States and weights for the $\mathbf{8}$ in $\mathbf{3} \otimes \bar{\mathbf{3}}$	
State	*Weight*
$u \otimes \bar{s}$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
$u \otimes \bar{d}$	$(1, 0)$
$d \otimes \bar{s}$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{2}}(d \otimes \bar{d} - u \otimes \bar{u})$, $\frac{1}{\sqrt{6}}(d \otimes \bar{d} + u \otimes \bar{u} - 2s \otimes \bar{s})$	$(0, 0)$
$d \otimes \bar{u}$	$(-1, 0)$
$s \otimes \bar{u}$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$s \otimes \bar{d}$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$

Removing these weights from the weight diagram, one is left with a singlet **1** with weight $(0,0)$, corresponding to the state

$$\frac{1}{\sqrt{3}}(u \otimes \bar{u} + s \otimes \bar{s} + d \otimes \bar{d}) \tag{4.39}$$

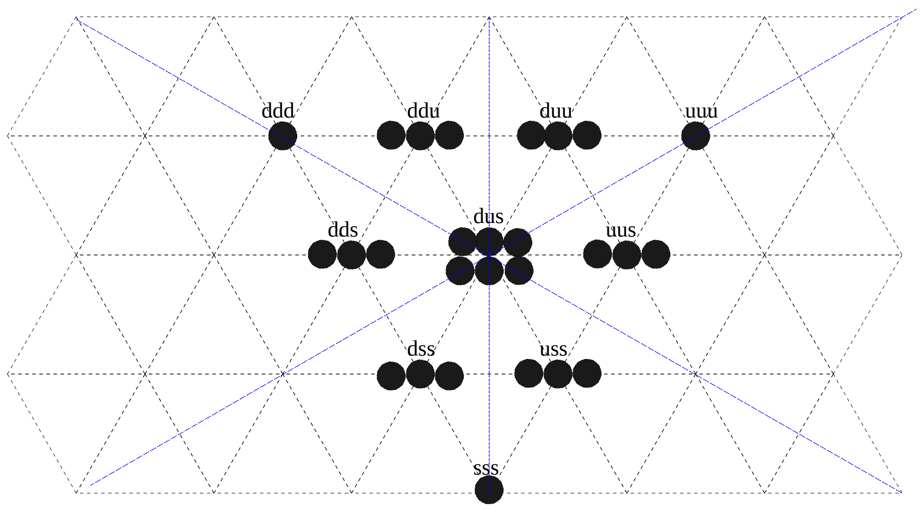
which is the unique linear combination- up to an overall scale- of $u \otimes \bar{u}$, $s \otimes \bar{s}$ and $d \otimes \bar{d}$ which is annihilated by the raising operators E_{+}^{m}. Hence we have the decomposition $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}$.

4.3.3 $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$ decomposition.

For this tensor product the quark content/weight table is as follows:

Quark content and weights for $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$	
Quark Content	*Weight*
$u \otimes u \otimes u$	$(\frac{3}{2}, \frac{\sqrt{3}}{2})$
$s \otimes s \otimes s$	$(0, -\sqrt{3})$
$d \otimes d \otimes d$	$(-\frac{3}{2}, \frac{\sqrt{3}}{2})$
$u \otimes u \otimes s$, $u \otimes s \otimes u$, $s \otimes u \otimes u$	$(1, 0)$
$u \otimes u \otimes d$, $u \otimes d \otimes u$, $d \otimes u \otimes u$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
$s \otimes s \otimes u$, $s \otimes u \otimes s$, $u \otimes s \otimes s$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$s \otimes s \otimes d$, $s \otimes d \otimes s$, $d \otimes s \otimes s$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$d \otimes d \otimes s$, $d \otimes s \otimes d$, $s \otimes d \otimes d$	$(-1, 0)$
$d \otimes d \otimes u$, $d \otimes u \otimes d$, $u \otimes d \otimes d$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$
$u \otimes d \otimes s$, $u \otimes s \otimes d$, $d \otimes u \otimes s$, $d \otimes s \otimes u$, $s \otimes u \otimes d$, $s \otimes d \otimes u$	$(0, 0)$

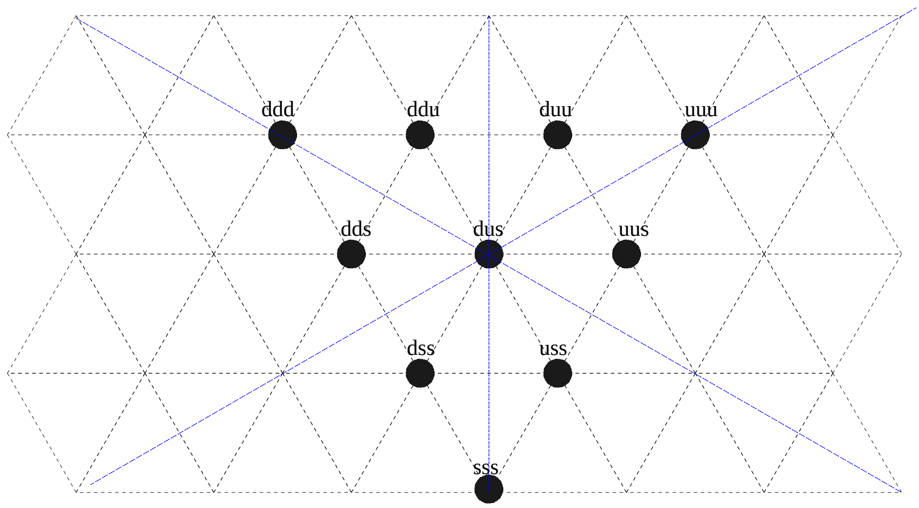
with weight diagram



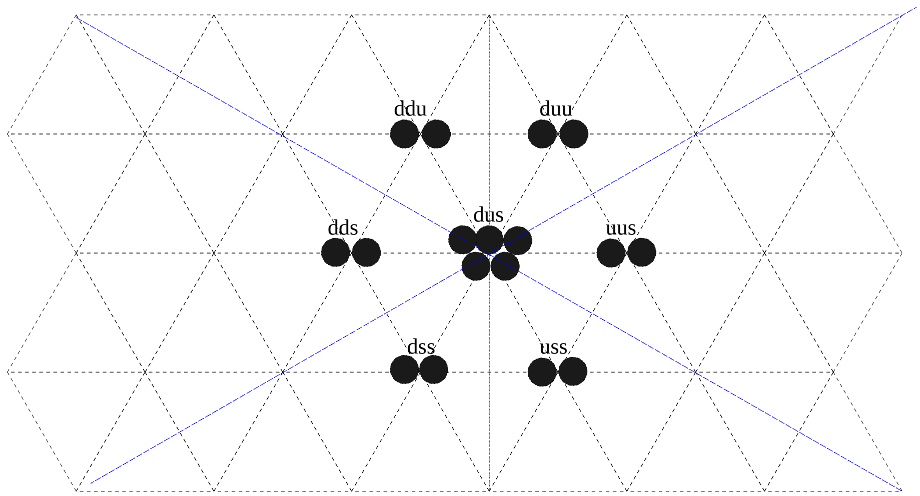
The raising and lowering operators are $E_{\pm}^{m} = e_{\pm}^{m} \otimes 1 \otimes 1 + 1 \otimes e_{\pm}^{m} \otimes 1 + 1 \otimes 1 \otimes e_{\pm}^{m}$. There are six weights of multiplicity 3, and the weight $(0,0)$ has multiplicity 6. The highest weight is $u \otimes u \otimes u$ with weight $(\frac{3}{2}, \frac{\sqrt{3}}{2})$. By applying lowering operators to this state, one obtains a triangular 10-dimensional irreducible representation denoted by **10**, which has normalized states and weights:

States and weights for $\mathbf{10}$ in $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$	
State	*Weight*
$u \otimes u \otimes u$	$(\frac{3}{2}, \frac{\sqrt{3}}{2})$
$s \otimes s \otimes s$	$(0, -\sqrt{3})$
$d \otimes d \otimes d$	$(-\frac{3}{2}, \frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{3}}(u \otimes u \otimes s + u \otimes s \otimes u + s \otimes u \otimes u)$	$(1, 0)$
$\frac{1}{\sqrt{3}}(u \otimes u \otimes d + u \otimes d \otimes u + d \otimes u \otimes u)$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{3}}(s \otimes s \otimes u + s \otimes u \otimes s + u \otimes s \otimes s)$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{3}}(s \otimes s \otimes d + s \otimes d \otimes s + d \otimes s \otimes s)$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{3}}(d \otimes d \otimes s + d \otimes s \otimes d + s \otimes d \otimes d)$	$(-1, 0)$
$\frac{1}{\sqrt{3}}(d \otimes d \otimes u + d \otimes u \otimes d + u \otimes d \otimes d)$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{6}}(u \otimes d \otimes s + u \otimes s \otimes d + d \otimes u \otimes s +$ $d \otimes s \otimes u + s \otimes u \otimes d + s \otimes d \otimes u)$	$(0, 0)$

The **10** weight diagram is



Removing the (non-vanishing) span of these states from the tensor product space, one is left with a 17-dimensional vector space. The new weight diagram is



Note that the highest weight is now $(\frac{1}{2}, \frac{\sqrt{3}}{2})$. This weight has has multiplicity 2. It should be noted that the subspace consisting of linear combinations of $d \otimes u \otimes u$, $u \otimes d \otimes u$ and $u \otimes u \otimes d$ which is annihilated by all raising operators E_+^m is two-dimensional and is spanned by the two orthogonal states $\frac{1}{\sqrt{6}}(d \otimes u \otimes u + u \otimes d \otimes u - 2u \otimes u \otimes d)$ and $\frac{1}{\sqrt{2}}(d \otimes u \otimes u - u \otimes d \otimes u)$. By acting on these two states with all possible lowering operators, one obtains two **8** representations whose states are mutually orthogonal.

The states and weights of these two **8** representations are summarized below:

States and weights for an **8** in $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$	
State	*Weight*
$\frac{1}{\sqrt{6}}(d \otimes u \otimes u + u \otimes d \otimes u - 2u \otimes u \otimes d)$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{6}}(s \otimes u \otimes u + u \otimes s \otimes u - 2u \otimes u \otimes s)$	$(1, 0)$
$\frac{1}{\sqrt{6}}(2d \otimes d \otimes u - d \otimes u \otimes d - u \otimes d \otimes d)$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$
$\frac{1}{2\sqrt{3}}(s \otimes d \otimes u + s \otimes u \otimes d + d \otimes s \otimes u$ $+u \otimes s \otimes d - 2d \otimes u \otimes s - 2u \otimes d \otimes s)$, $\frac{1}{2\sqrt{3}}(2s \otimes d \otimes u + 2d \otimes s \otimes u - s \otimes u \otimes d$ $-d \otimes u \otimes s - u \otimes s \otimes d - u \otimes d \otimes s)$	$(0, 0)$
$\frac{1}{\sqrt{6}}(s \otimes d \otimes d + d \otimes s \otimes d - 2d \otimes d \otimes s)$	$(-1, 0)$
$\frac{1}{\sqrt{6}}(2s \otimes s \otimes u - s \otimes u \otimes s - u \otimes s \otimes s)$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{6}}(2s \otimes s \otimes d - s \otimes d \otimes s - d \otimes s \otimes s)$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$

States and weights for another **8** in $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$	
State	*Weight*
$\frac{1}{\sqrt{2}}(d \otimes u \otimes u - u \otimes d \otimes u)$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{2}}(s \otimes u \otimes u - u \otimes s \otimes u)$	$(1, 0)$
$\frac{1}{\sqrt{2}}(d \otimes u \otimes d - u \otimes d \otimes d)$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$
$\frac{1}{2}(s \otimes d \otimes u + s \otimes u \otimes d - d \otimes s \otimes u - u \otimes s \otimes d)$, $\frac{1}{2}(s \otimes u \otimes d + d \otimes u \otimes s - u \otimes s \otimes d - u \otimes d \otimes s)$	$(0, 0)$
$\frac{1}{\sqrt{2}}(s \otimes d \otimes d - d \otimes s \otimes d)$	$(-1, 0)$
$\frac{1}{\sqrt{2}}(s \otimes u \otimes s - u \otimes s \otimes s)$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$
$\frac{1}{\sqrt{2}}(s \otimes d \otimes s - d \otimes s \otimes s)$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$

Removing these weights from the weight diagram, we are left with a singlet **1** with weight $(0, 0)$. The state corresponding to this singlet is

$$\frac{1}{\sqrt{6}}(s \otimes d \otimes u - s \otimes u \otimes d + d \otimes u \otimes s - d \otimes s \otimes u + u \otimes s \otimes d - u \otimes d \otimes s) \tag{4.40}$$

which is the only linear combination-up to overall scale- of $u \otimes d \otimes s$, $u \otimes s \otimes d$, $d \otimes u \otimes s$, $d \otimes s \otimes u$, $s \otimes u \otimes d$ and $s \otimes d \otimes u$ which is annihilated by all the raising operators.

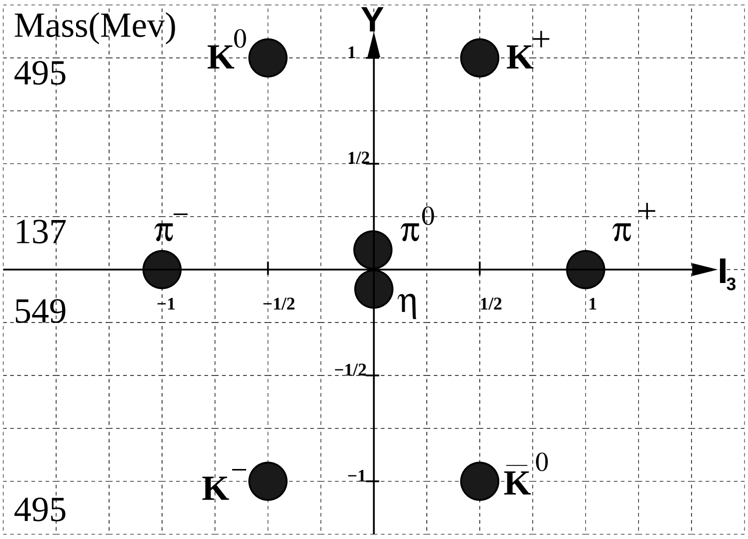
Hence we have the decomposition $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$ where the states in **10** are symmetric, but the state in **1** is antisymmetric. The **8** states have mixed symmetry.

4.4 The Quark Model

It is possible to arrange the baryons and the mesons into $SU(3)$ multiplets; i.e. the states lie in Hilbert spaces which are tensor products of vector spaces equipped with irreducible representations of $\mathcal{L}(SU(3))$. To see examples of this, it is convenient to group hadrons into multiplets with the same baryon number and spin. We plot the hypercharge $Y = S + B$ where S is the strangeness and B is the baryon number against the isospin eigenvalue I_3 for these particles.

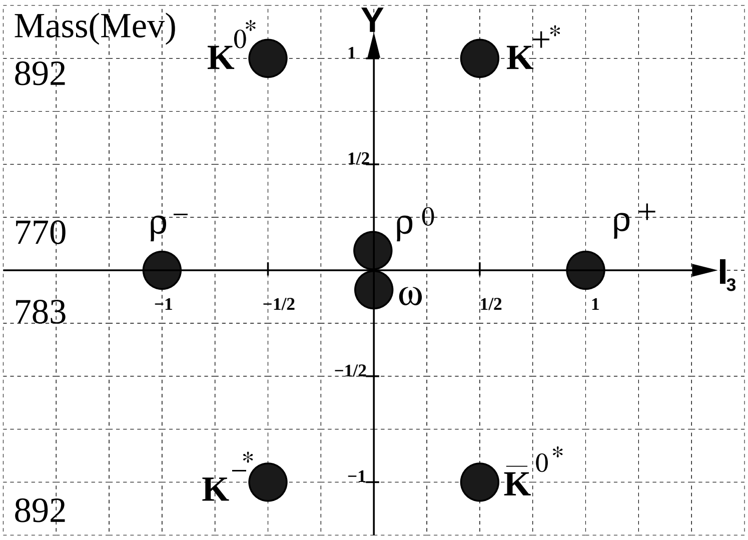
4.4.1 Meson Multiplets

The pseudoscalar meson octet has $B = 0$ and $J = 0$. The (I_3, Y) diagram is



There is also a $J = 0$ meson singlet η'.

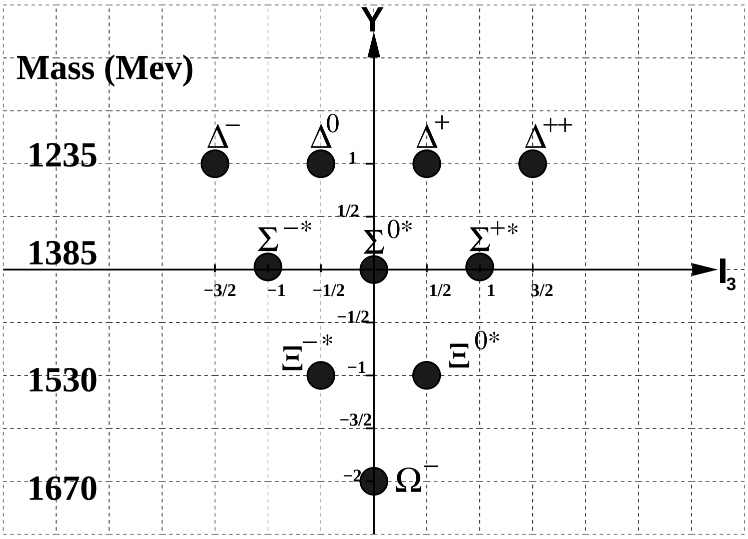
The vector meson octet has $B = 0$ and $J = 1$. The (I_3, Y) diagram is



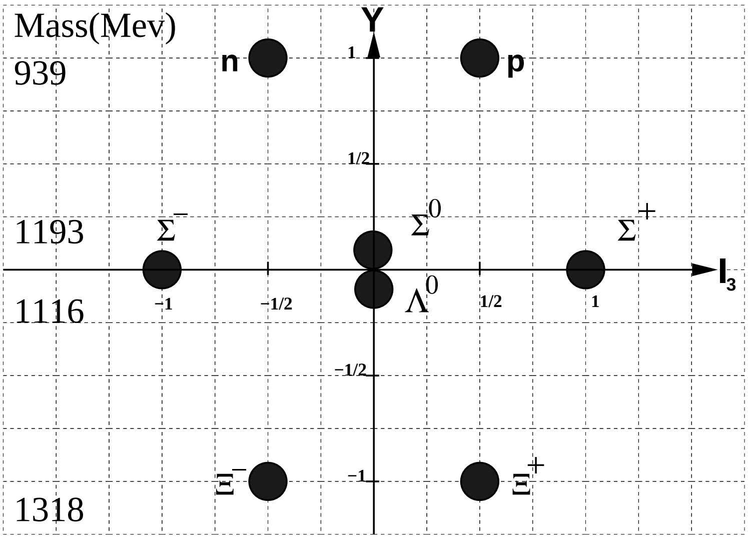
There is also a $J = 1$ meson singlet, ϕ.

4.4.2 Baryon Multiplets

The baryon decuplet has $B = 1$ and $J = \frac{3}{2}$ with (I_3, Y) diagram



There is also an antibaryon decuplet with $(I_3, Y) \to -(I_3, Y)$.
The baryon octet has $B = 1$, $J = \frac{1}{2}$ with (I_3, Y) diagram



and there is also a $J = \frac{1}{2}$ baryon singlet Λ^{0*}.

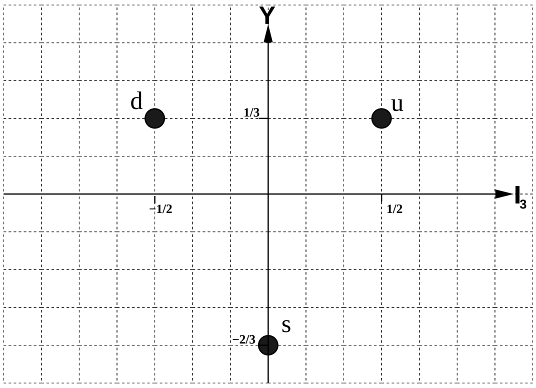
4.4.3 Quarks: Flavour and Colour

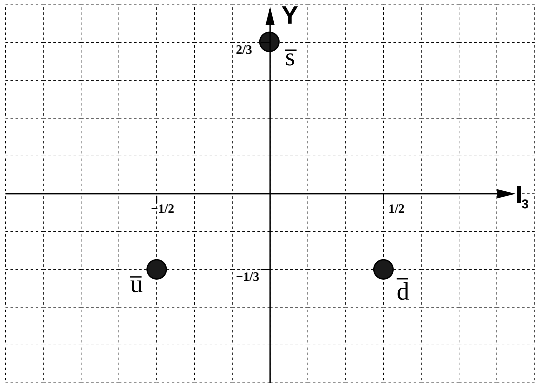
On making the identification $(p, q) = (I_3, \frac{\sqrt{3}}{2} Y)$ the points on the meson and baryon octets and the baryon decuplet can be matched to points on the weight diagrams of the **8** and **10** of $\mathcal{L}(SU(3))$.

Motivated by this, it is consistent to consider the (light) meson states as lying within a $\mathbf{3} \otimes \bar{\mathbf{3}}$; as $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}$, the meson octets are taken to correspond to the **8** states, and the meson singlets correspond to the singlet **1** states. The light baryon states lie within a $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$; the baryon decuplet corresponds to the **10** in $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$; the baryon octet corresponds to appropriate linear combinations of elements in the **8** irreps, and the baryon singlet corresponds to the **1**.

In this model, the fundamental states in the **3** are quarks, with basis states u (up), d (down) and s (strange). The basis labels u, d, s are referred to as the *flavours* of the quarks. The $\bar{\mathbf{3}}$ states are called antiquarks with basis $\bar{u}$, $\bar{d}$, $\bar{s}$. Baryons are composed of bound states of three quarks qqq, mesons are composed of bound states of pairs of quarks and antiquarks $q\bar{q}$. The quarks have $J = \frac{1}{2}$ and $B = \frac{1}{3}$ whereas the antiquarks have $J = \frac{1}{2}$ and $B = -\frac{1}{3}$ which is consistent with the values of B and J for the baryons and mesons.

The quark and antiquark flavours can be plotted on the (I_3, Y) plane:





We have shown that mesons and baryons can be constructed from $q\bar{q}$ and qqq states respectively. But why do qq particles not exist? This problem is resolved using the notion of colour. Consider the Δ^{++} particle in the baryon decuplet. This is a $u \otimes u \otimes u$ state with $J = \frac{3}{2}$. The members of the decuplet are the spin $\frac{3}{2}$ baryons of lowest mass, so we assume that the quarks have vanishing orbital angular momentum. Then the spin $J = \frac{3}{2}$ is obtained by having all the quarks in the spin up state, i.e. $u \uparrow \otimes u \uparrow \otimes u \uparrow$. However, this violates the Pauli exclusion principle. To get round this problem, it is conjectured that quarks possess additional labels other than flavour. In particular, quarks have additional charges called *colour* charges- there are three colour basis states associated with quarks called r (red), g (green) and b (blue). The quark state wave-functions contain colour factors which lie in a **3** representation of $SU(3)$ which describes their colour; the colour of antiquark states corresponds to a $\bar{\mathbf{3}}$ representation of $SU(3)$ (colour). This colour $SU(3)$ is independent of the flavour $SU(3)$.

These colour charges are also required to remove certain discrepancies (of powers of 3) between experimentally observed processes such as the decay $\pi^0 \to 2\gamma$ and the cross section ratio between the processes $e^+e^- \to$ hadrons and $e^+e^- \to \mu^+\mu^-$ and theoretical predictions. However, although colour plays an important role in these processes, it seems that one cannot measure colour directly experimentally- all known mesons and baryons are $SU(3)$ colour singlets (so colour is confined). This principle excludes the possibility of having qq particles, as there is no singlet state in the $SU(3)$ (colour) tensor product decomposition $\mathbf{3} \otimes \mathbf{3}$, though there is in $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$ and $\mathbf{3} \otimes \bar{\mathbf{3}}$. Other products of **3** and $\bar{\mathbf{3}}$ can also be ruled out in this fashion.

Nevertheless, the decomposition of $\mathbf{3} \otimes \mathbf{3}$ is useful because it is known that in addition to the u, d and s quark states, there are also c (charmed), t (top) and b (bottom) quark flavours. However, the c, t and b quarks are heavier than the u, d and s quarks, and are unstable- they decay into the lighter quarks. The $SU(3)$ symmetry cannot be meaningfully extended to a naive $SU(6)$ symmetry because of the large mass differences which break the symmetry. In this context, meson states formed from a heavy antiquark and a light quark can only be reliably put into $\mathbf{3}$ multiplets, whereas baryons made from one heavy and two light quarks lie in $\mathbf{3} \otimes \mathbf{3} = 6 \oplus \bar{\mathbf{3}}$ multiplets.

5. Spacetime Symmetry

In this section we examine spacetime symmetry in the absence of gravity. Spacetime is taken to be 4-dimensional Minkowski space, M^4 , with real co-ordinates x^μ for $\mu = 0, 1, 2, 3$, equipped with the Minkowski metric which has the non-vanishing components $\eta_{00} = -\eta_{11} = -\eta_{22} = -\eta_{33} = 1$, or as a matrix

$$(\eta)_{\mu\nu} = (\eta)^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{5.1}$$

5.1 The Lorentz Group

A Lorentz transformation is a linear transformation $\Lambda : M^4 \to M^4$ which transforms co-ordinates

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{5.2}$$

for $\Lambda^\mu{}_\nu \in \mathbb{R}$, but leaves the length invariant

$$\eta_{\mu\nu} x'^\mu x'^\nu = \eta_{\mu\nu} x^\mu x^\nu \tag{5.3}$$

for all x.

This condition can be rewritten in matrix notation as

$$\Lambda^T \eta \Lambda = \eta \tag{5.4}$$

Suppose that Λ_1, Λ_2 are two 4×4 matrices satisfying (5.4) then it is straightforward to see that $\Lambda_1 \Lambda_2$ satisfies (5.4). Also, if Λ satisfies (5.4) then $\det \Lambda = \pm 1$, hence Λ is invertible, with inverse $\Lambda^{-1} = \eta^{-1} \Lambda^T \eta$, and

$$\begin{aligned} (\Lambda^{-1})^T \eta \Lambda^{-1} &= \eta \Lambda \eta^{-1} . \eta . \eta^{-1} \Lambda^T \eta \\ &= \eta \Lambda \eta^{-1} \Lambda^T \eta \\ &= \eta \Lambda \Lambda^{-1} \\ &= \eta \end{aligned} \tag{5.5}$$

so Λ^{-1} is also a Lorentz transformation. Hence, the set of Lorentz transformations forms a group, under matrix multiplication.

Write a generic Lorentz transformation as

$$\Lambda = \begin{pmatrix} \lambda & \underline{\beta}^T \\ \underline{\alpha} & R \end{pmatrix} \tag{5.6}$$

where $\lambda \in \mathbb{R}$, $\underline{\alpha}, \underline{\beta} \in \mathbb{R}^3$ and R is a 3×3 real matrix

Then the constraint (5.4) is equivalent to

$$\lambda^2 = 1 + \underline{\alpha} . \underline{\alpha} \tag{5.7}$$

and

$$\lambda \underline{\beta} = R^T \underline{\alpha} \tag{5.8}$$

and

$$R^T R - \underline{\beta} \underline{\beta}^T = \mathbb{I}_3 \tag{5.9}$$

Note that (5.7) implies that

$$\lambda = \pm \sqrt{1 + \underline{\alpha}.\underline{\alpha}} \tag{5.10}$$

so in particular $\lambda \leq -1$ or $\lambda \geq +1$. Then (5.8) fixes $\underline{\beta}$ in terms of $\underline{\alpha}$ and R by

$$\underline{\beta} = \pm \frac{1}{\sqrt{1 + \underline{\alpha}.\underline{\alpha}}} R^T \underline{\alpha} \tag{5.11}$$

which can be used to rewrite (5.9) as

$$R^T R - \frac{1}{1 + \underline{\alpha}.\underline{\alpha}} R^T \underline{\alpha} \underline{\alpha}^T R = \mathbb{I}_3 \tag{5.12}$$

Define

$$\hat{R} = \big(1 - \frac{1}{\sqrt{1 + \underline{\alpha}.\underline{\alpha}}(1 + \sqrt{1 + \underline{\alpha}.\underline{\alpha}})} \underline{\alpha} \underline{\alpha}^T \big) R \tag{5.13}$$

or equivalently

$$R = \big(1 + \frac{1}{1 + \sqrt{1 + \underline{\alpha}.\underline{\alpha}}} \underline{\alpha} \underline{\alpha}^T \big) \hat{R} \tag{5.14}$$

Then (5.12) implies

$$\hat{R}^T \hat{R} = \mathbb{I}_3 \tag{5.15}$$

i.e. $\hat{R} \in O(3)$. Moreover, it is straightforward to check directly that

$$\det \hat{R} = \frac{1}{\sqrt{1 + \underline{\alpha}.\underline{\alpha}}} \det R \tag{5.16}$$

where we have used the formula $\det(\mathbb{I}_3 + K \underline{\alpha} \underline{\alpha}^T) = 1 + K \underline{\alpha}.\underline{\alpha}$ for any K. Also, one can write

$$\Lambda = \begin{pmatrix} \lambda & 0 \\ \underline{\alpha} & \mathbb{I}_3 \end{pmatrix} \begin{pmatrix} 1 & \lambda^{-2} \underline{\alpha}^T \\ 0 & \mathbb{I}_3 - \lambda^{-2} \underline{\alpha} \underline{\alpha}^T \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \tag{5.17}$$

and hence

$$\det \Lambda = \frac{\lambda}{1 + \underline{\alpha}.\underline{\alpha}} \det R = \frac{\lambda}{\sqrt{1 + \underline{\alpha}.\underline{\alpha}}} \det \hat{R} \tag{5.18}$$

$O(3)$ has two connected components, the connected component of $\mathbb{I}_3$ (which is $SO(3)$) whose elements have determinant $\det \hat{R} = +1$, and the connected component of $-\mathbb{I}_3$, whose elements have determinant $\det \hat{R} = -1$.

There are therefore four connected components of the Lorentz group, according as ${\Lambda^0}_0 \leq -1$ or ${\Lambda^0}_0 \geq 1$ and $\det \Lambda = +1$ or $\det \Lambda = -1$. It is not possible to construct a smooth curve in the Lorentz group passing from one of these components to the other.

The set of Lorentz transformations with $\det \Lambda = +1$ forms a subgroup of the Lorentz group.

Note that (5.13) implies that

$$\hat{R}^T \underline{\alpha} = \frac{1}{\sqrt{1+\underline{\alpha}.\underline{\alpha}}} R^T \underline{\alpha} \tag{5.19}$$

and hence

$$\underline{\beta}.\underline{\beta} = \frac{1}{1+\underline{\alpha}.\underline{\alpha}} \underline{\alpha}^T R R^T \underline{\alpha} = \underline{\alpha}.\underline{\alpha} \tag{5.20}$$

So, if Λ and Λ' are two Lorentz transformations

$$\Lambda = \begin{pmatrix} \lambda & \underline{\beta}^T \\ \underline{\alpha} & R \end{pmatrix} \qquad \Lambda' = \begin{pmatrix} \lambda' & \underline{\beta}'^T \\ \underline{\alpha}' & R' \end{pmatrix} \tag{5.21}$$

with $\lambda \geq 1$ and $\lambda' \geq 1$, then

$$(\Lambda\Lambda')^0{}_0 = \lambda\lambda' + \underline{\beta}.\underline{\alpha}' \geq \sqrt{1+\underline{\alpha}.\underline{\alpha}}\sqrt{1+\underline{\alpha}'.\underline{\alpha}'} - \sqrt{\underline{\alpha}.\underline{\alpha}}\sqrt{\underline{\alpha}'.\underline{\alpha}'} \geq 1 \tag{5.22}$$

Hence the set of Lorentz transformations with $\Lambda^0{}_0 \geq 1$ also forms a subgroup of the Lorentz group.

The subgroup of Lorentz transformations with $\det\Lambda = +1$ and $\Lambda^0{}_0 \geq 1$ is called the proper orthochronous Lorentz group, which we denote by $SO(3,1)^{\uparrow}$.

We note the useful lemma

Lemma 10. *Suppose that $\Lambda \in SO(3,1)^{\uparrow}$. Then there exist S_1, $S_2 \in SO(3)$ and $z \in \mathbb{R}$ such that*

$$\Lambda = \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & S_1 \end{pmatrix} \begin{pmatrix} \cosh z & \sinh z & 0 & 0 \\ \sinh z & \cosh z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & S_2 \end{pmatrix} \tag{5.23}$$

Proof

From the analysis of the Lorentz group so far, we have shown that if $\Lambda \in SO(3,1)^{\uparrow}$ then there exists $\underline{\alpha} \in \mathbb{R}^3$ and $\hat{R} \in SO(3)$ such that

$$\begin{aligned} \Lambda &= \begin{pmatrix} \sqrt{1+\underline{\alpha}.\underline{\alpha}} & \underline{\alpha}^T \hat{R} \\ \underline{\alpha} & (1 + \frac{1}{1+\sqrt{1+\underline{\alpha}.\underline{\alpha}}}\underline{\alpha\alpha}^T)\hat{R} \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{1+\underline{\alpha}.\underline{\alpha}} & \underline{\alpha}^T \\ \underline{\alpha} & 1 + \frac{1}{1+\sqrt{1+\underline{\alpha}.\underline{\alpha}}}\underline{\alpha\alpha}^T \end{pmatrix} \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \hat{R} \end{pmatrix} \end{aligned} \tag{5.24}$$

There also exists $S_1 \in SO(3)$ and $z \in \mathbb{R}$ such that

$$\underline{\alpha} = S_1 \begin{pmatrix} \sinh z \\ 0 \\ 0 \end{pmatrix} \tag{5.25}$$

The result follows on substituting this into Λ and setting $S_2 = (S_1)^T \hat{R}$. ∎

5.2 The Lorentz Group and $SL(2,\mathbb{C})$

Consider the spacetime co-ordinates x^μ. Define the matrices

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.26}$$

Then given real spacetime co-ordinates x^μ, define the 2×2 complex hermitian matrix

$$\tilde{x} = x_\mu \sigma^\mu = \begin{pmatrix} x^0 - x^3 & -x^1 + ix^2 \\ -x^1 - ix^2 & x^0 + x^3 \end{pmatrix} \tag{5.27}$$

Observe that *any* hermitian 2×2 matrix can be written as $\tilde{x}$ for some real x^μ.

Note that

$$\det \tilde{x} = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = \eta_{\mu\nu} x^\mu x^\nu \tag{5.28}$$

$\eta_{\mu\nu} x^\mu x^\nu$ is invariant under the action of $SO(3,1)$. $\det \tilde{x}$ is invariant under the action of $SL(2,\mathbb{C})$, the complex 2×2 matrices with unit determinant.

Proposition 30. *There exists an isomorphism $\pi : SL(2,\mathbb{C})/\mathbb{Z}_2 \to SO(3,1)^\uparrow$ where $SL(2,\mathbb{C})/\mathbb{Z}_2$ consists of elements $\pm N \in SL(2,\mathbb{C})$ with $+N$ identified with $-N$.*

Proof

Given $N \in SL(2,\mathbb{C})$ consider the 2×2 complex matrix $N\tilde{x}N^\dagger$. The components of this matrix are linear in the spacetime co-ordinates x^μ. As $\tilde{x}$ is hermitian, it follows that $N\tilde{x}N^\dagger$ is also hermitian. Hence there exist $\Lambda^\mu{}_\nu \in \mathbb{R}$ (independent of x) for $\mu, \nu = 0, \ldots, 3$ such that

$$N\tilde{x}N^\dagger = \widetilde{(\Lambda x)} \tag{5.29}$$

Taking the determinant of both sides we find $\det \tilde{x} = \det \widetilde{(\Lambda x)}$ for all x, and therefore Λ is a Lorentz transformation.

Set

$$\Lambda = \pi(N) \tag{5.30}$$

Note that

$$\text{Tr } \widetilde{(\Lambda x)} = 2\Lambda^0{}_\mu x^\mu = \text{Tr } (N^\dagger N \tilde{x}) \tag{5.31}$$

Setting $x^0 = 1$, $x^1 = x^2 = x^3 = 0$ we find $\Lambda^0{}_0 = \frac{1}{2}\text{Tr } (N^\dagger N) > 0$.

If $N_1, N_2 \in SL(2,\mathbb{C})$ then

$$(N_1 N_2)\tilde{x}(N_1 N_2)^\dagger = (\widetilde{\pi(N_1 N_2) x}) \tag{5.32}$$

But

$$\begin{aligned} (N_1 N_2)\tilde{x}(N_1 N_2)^\dagger &= N_1 (N_2 \tilde{x} N_2^\dagger) N_1^\dagger \\ &= N_1 (\widetilde{\pi(N_2) x}) N_1^\dagger \\ &= \widetilde{\pi(N_1) \pi(N_2) x} \end{aligned} \tag{5.33}$$

Hence $(\widetilde{\pi(N_1 N_2)}x) = \widetilde{\pi(N_1)\pi(N_2)}x$ for all x, which implies $\pi(N_1 N_2) = \pi(N_1)\pi(N_2)$.

Next we will establish that π is onto $SO(3,1)^{\uparrow}$. First recall that any $R \in SO(3,1)^{\uparrow}$ of the form

$$R = \begin{pmatrix} 1 & \underline{0}^T \\ \underline{0} & \hat{R} \end{pmatrix} \tag{5.34}$$

can be written as a product of rotations around the spatial co-ordinate axes

$$R_1(\phi_1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\phi_1 & \sin\phi_1 \\ 0 & 0 & -\sin\phi_1 & \cos\phi_1 \end{pmatrix} \tag{5.35}$$

$$R_2(\phi_2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi_2 & 0 & -\sin\phi_2 \\ 0 & 0 & 1 & 0 \\ 0 & \sin\phi_2 & 0 & \cos\phi_2 \end{pmatrix} \tag{5.36}$$

and

$$R_3(\phi_3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi_3 & \sin\phi_3 & 0 \\ 0 & -\sin\phi_3 & \cos\phi_3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.37}$$

By a direct computation we find $\pi(e^{\frac{i\phi_j}{2}\sigma^j}) = R_j$ for $j = 1,2,3$; and

$$\pi(e^{-\frac{z}{2}\sigma^1}) = \begin{pmatrix} \cosh z & \sinh z & 0 & 0 \\ \sinh z & \cosh z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.38}$$

where $e^{\frac{i\phi_j}{2}\sigma^j} \in SL(2,\mathbb{C})$ for $j = 1,2,3$ and $e^{-\frac{z}{2}\sigma^1} \in SL(2,\mathbb{C})$.

Hence, if $\Lambda \in SO(3,1)^{\uparrow}$, it follows that one can write $\Lambda = \Lambda_1.\Lambda_2 \dots \Lambda_k$ where Λ_i are elementary rotation or boost transformations in $SO(3,1)^{\uparrow}$, and from the above reasoning, $\Lambda_i = \pi(N_i)$ for some $N_i \in SL(2,\mathbb{C})$. Therefore, $\Lambda = \pi(N_1.N_2 \dots N_k)$, so π is onto $SO(3,1)^{\uparrow}$.

Next, suppose that $\pi(N) = \pi(M)$ for $N, M \in SL(2,\mathbb{C})$. Then

$$N\tilde{x}N^{\dagger} = M\tilde{x}M^{\dagger} \tag{5.39}$$

Set $Q = M^{-1}N$, so that

$$Q\tilde{x}Q^{\dagger} = \tilde{x} \tag{5.40}$$

Setting $x^0 = 1$, $x^1 = x^2 = x^3 = 0$, we obtain $QQ^{\dagger} = \mathbb{I}_2$, so $Q \in SU(2)$. Hence

$$Q\sigma^i = \sigma^i Q \tag{5.41}$$

for $i = 1, 2, 3$. The only $Q \in SU(2)$ satisfying this is $Q = \pm\mathbb{I}_2$, so $M = \pm N$. Hence π is a $2-1$ map.

Lastly, we must prove that if $N \in SL(2,\mathbb{C})$ then $\pi(N) \in SO(3,1)^{\uparrow}$. We have already shown that $\pi(N)$ is orthochronous. Suppose that $\det(\pi(N)) = -1$. Consider

$$\hat{\Lambda} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \pi(N) \tag{5.42}$$

The $\det \hat{\Lambda} = +1$, so $\hat{\Lambda} \in SO(3,1)^{\uparrow}$. Hence, there exists some $N' \in SL(2,\mathbb{C})$ such that $\hat{\Lambda} = \pi(N')$, so

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \pi(N) = \pi(N') \tag{5.43}$$

Setting $Y = N'N^{-1}$, we obtain

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \pi(Y) \tag{5.44}$$

for some $Y \in SL(2,\mathbb{C})$. This implies that

$$Y x_\mu \sigma^\mu Y^\dagger = x_0 \mathbb{I}_2 + x_1 \sigma^1 + x_2 \sigma^2 - x_3 \sigma^3 \tag{5.45}$$

for all x_μ. In particular, for $x^0 = 1$, $x^1 = x^2 = x^3 = 0$ we find $YY^\dagger = \mathbb{I}_2$, so $Y \in SU(2)$. The remaining constraints are

$$Y\sigma^1 = \sigma^1 Y, \quad Y\sigma^2 = \sigma^2 Y, \quad Y\sigma^3 = -\sigma^3 Y \tag{5.46}$$

This is not possible, because $[Y, \sigma^1] = [Y, \sigma^2] = 0$ implies that $Y = \alpha\mathbb{I}_2$ for some $\alpha \in \mathbb{C}$. As $\det Y = 1$ this implies $Y = \pm\mathbb{I}_2$, but then $Y\sigma^3 \neq -\sigma^3 Y$. Hence if $N \in SL(2,\mathbb{C})$ then $\pi(N) \in SO(3,1)^{\uparrow}$.

Although $\pi : SL(2,\mathbb{C}) \to SO(3,1)^{\uparrow}$ is not 1-1, we have shown that the restriction of π to $SL(2,\mathbb{C})/\mathbb{Z}_2$, in which N is identified with $-N$ is 1-1. ■

5.3 The Lie Algebra $\mathcal{L}(SO(3,1))$

To compute the constraints on the tangent matrices, consider a curve in the Lorentz group $\Lambda(t)$ with $\Lambda(0) = \mathbb{I}_4$. This is constrained by

$$\Lambda(t)^T \eta \Lambda(t) = \eta \tag{5.47}$$

Differentiate this constraint and set $t = 0$ to obtain

$$m^T \eta + \eta m = 0 \tag{5.48}$$

where $m = (\frac{d\Lambda(t)}{dt})|_{t=0}$. The generic solution to this constraint is

$$m^\mu{}_\nu = \begin{pmatrix} 0 & \underline{\chi} \\ \underline{\chi}^T & S \end{pmatrix} \tag{5.49}$$

for any $\underline{\chi} \in \mathbb{R}^3$ and S is a real 3×3 antisymmetric matrix; $S = -S^T$. There are three real degrees of freedom in $\underline{\chi}$ and three real degrees of freedom in the antisymmetric matrix S. Hence the Lie algebra is six-dimensional.

Define the 4×4 matrices $M^{\mu\nu}$ for $\mu, \nu = 0, 1, 2, 3$ by

$$(M^{\mu\nu})^\alpha{}_\beta = i(\eta^{\mu\alpha}\delta^\nu{}_\beta - \eta^{\nu\alpha}\delta^\mu{}_\beta) \tag{5.50}$$

note that $M^{\mu\nu} = -M^{\nu\mu}$, so there are only six linearly independent matrices defined here. By direct computation, we find

$$M^{01} = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad M^{02} = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
$$M^{03} = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \quad M^{12} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
$$M^{13} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \quad M^{23} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \tag{5.51}$$

and

$$[M^{\mu\nu}, M^{\rho\sigma}] = i\big(M^{\mu\sigma}\eta^{\nu\rho} + M^{\nu\rho}\eta^{\mu\sigma} - M^{\mu\rho}\eta^{\nu\sigma} - M^{\nu\sigma}\eta^{\mu\rho}\big) \tag{5.52}$$

which defines the complexified Lie algebra of the Lorentz group.

Define

$$\begin{aligned} J_i &= \frac{1}{2}\epsilon_{ijk}M_{jk} \\ K_i &= M_{0i} \end{aligned} \tag{5.53}$$

for $i, j, k = 1, 2, 3$. Then it follows that

$$\begin{aligned} [J_i, J_j] &= i\epsilon_{ijk}J_k \\ [K_i, K_j] &= -i\epsilon_{ijk}J_k \\ [J_i, K_j] &= i\epsilon_{ijk}K_k \end{aligned} \tag{5.54}$$

So, setting

$$A_i = \frac{1}{2}(J_i - iK_i), \qquad B_i = \frac{1}{2}(J_i + iK_i) \tag{5.55}$$

we obtain the commutation relations

$$\begin{aligned}
[A_i, A_j] &= i\epsilon_{ijk} A_k \\
[B_i, B_j] &= i\epsilon_{ijk} B_k \\
[A_i, B_j] &= 0
\end{aligned} \tag{5.56}$$

Hence the complexified Lorentz algebra $\mathcal{L}(SO(3,1))$ can be written as the direct sum of two commuting complexified $\mathcal{L}(SU(2))$ algebras. It follows that one can classify irreducible representations of the Lorentz algebra by spins (A, B) for $2A, 2B \in \mathbb{N}$.

5.4 Spinors and Invariant Tensors of $SL(2,\mathbb{C})$

Definition 44. *The left handed Weyl spinors are elements of a 2-dimensional complex vector space V on which the fundamental representation of $SL(2,\mathbb{C})$ acts via $\mathcal{D}(N)\psi = N\psi$ where $N \in SL(2,\mathbb{C})$. In terms of components, if $\psi \in V$ has components ψ_α for $\alpha = 1, 2$ with respect to some basis of V, then under the action of $SL(2,\mathbb{C})$, ψ transforms as*

$$\psi_\alpha \to \psi'_\alpha = N_\alpha{}^\beta \psi_\beta \tag{5.57}$$

where $N \in SL(2,\mathbb{C})$.

Definition 45. *The right handed Weyl spinors are elements of a 2-dimensional complex vector space $\bar{V}$ on which the complex conjugate of the fundamental representation of $SL(2,\mathbb{C})$ acts as $\mathcal{D}^*(N)\bar{\chi} = N^\star \bar{\chi}$ where $N \in SL(2,\mathbb{C})$ and $N^\star$ is the complex conjugate of N. In terms of components, if $\bar{\chi} \in \bar{V}$ has components $\bar{\chi}_{\dot{\alpha}}$ for $\dot{\alpha} = 1, 2$, then under the action of $SL(2,\mathbb{C})$, $\bar{\chi}$ transforms as*

$$\bar{\chi}_{\dot{\alpha}} \to \bar{\chi}'_{\dot{\alpha}} = N^*{}_{\dot{\alpha}}{}^{\dot{\beta}} \bar{\chi}_{\dot{\beta}} \tag{5.58}$$

where $N \in SL(2,\mathbb{C})$.

Note: One should regard α and $\dot{\alpha}$ as being entirely independent! The components of these spinors anticommute.

We also define $\epsilon^{\alpha\beta}$ and $\epsilon_{\alpha\beta}$ to be totally skew-symmetric with

$$\epsilon^{\alpha\beta} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \epsilon_{\alpha\beta} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{5.59}$$

and observe that $\epsilon^{\alpha\beta}\epsilon_{\beta\gamma} = \delta^\alpha{}_\gamma$. One defines $\epsilon^{\dot{\alpha}\dot{\beta}}$ and $\epsilon_{\dot{\alpha}\dot{\beta}}$ similarly.

Note that $\epsilon_{\alpha\beta}$ is invariant under $SL(2,\mathbb{C})$, as

$$\epsilon_{\alpha\beta} \to \epsilon'_{\alpha\beta} = \epsilon_{\mu\nu} N_\alpha{}^\mu N_\beta{}^\nu = \det(N)\epsilon_{\alpha\beta} = \epsilon_{\alpha\beta} \tag{5.60}$$

or in matrix notation $N\epsilon N^T = \epsilon$.

If we define the *contravariant representation* $\mathcal{D}_{CV}$ on V via

$$\mathcal{D}_{CV}(N)\psi \to (N^T)^{-1}\psi \tag{5.61}$$

and the complex conjugate contravariant representation by

$$\mathcal{D}^*_{CV}(N)\bar{\chi} \to (N^{*T})^{-1}\bar{\chi} \tag{5.62}$$

then $N\epsilon N^T = \epsilon$ implies that $(N^T)^{-1} = \epsilon^{-1}N\epsilon$, so $\mathcal{D}_{CV}$ is equivalent to the fundamental representation. The complex conjugate representations are similarly equivalent.

The tensors $\epsilon_{\alpha\beta}$ and $\epsilon^{\alpha\beta}$ are called invariant tensors as they transform into themselves under the action of $SL(2,\mathbb{C})$. For $SO(3,1)$, the invariant tensors are $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$, which can be used to raise and lower indices. We will raise and lower $SL(2,\mathbb{C})$ indices using $\epsilon^{\alpha\beta}$, $\epsilon_{\alpha\beta}$, so if ψ_α, $\bar{\chi}_{\dot{\alpha}}$ are in the fundamental and conjugate representations respectively, then we define

$$\psi^\alpha = \epsilon^{\alpha\beta}\psi_\beta, \qquad \bar{\chi}^{\dot{\alpha}} = \epsilon^{\dot{\alpha}\dot{\beta}}\bar{\chi}_{\dot{\beta}} \tag{5.63}$$

One can construct a tensor product representation of the fundamental representation acting on n products of V, $V \otimes V \otimes V \cdots \otimes V$. In terms of components, elements of the tensor product vector space have components $\psi_{\alpha_1 \ldots \alpha_n}$ which transform under the action of $SL(2,\mathbb{C})$ as

$$\psi_{\alpha_1,\ldots,\alpha_n} \to \psi'_{\alpha_1,\ldots,\alpha_n} = N_{\alpha_1}{}^{\beta_1} \ldots N_{\alpha_n}{}^{\beta_n}\psi_{\beta_1\ldots\beta_n} \tag{5.64}$$

for $N \in SL(2,\mathbb{C})$. Similarly, tensor product representations of the complex conjugate representation correspond to complex tensors with components $\bar{\chi}_{\dot{\alpha}_1\ldots\dot{\alpha}_m}$ which transform as

$$\bar{\chi}_{\dot{\alpha}_1\ldots\dot{\alpha}_m} \to \bar{\chi}'_{\dot{\alpha}_1\ldots\dot{\alpha}_m} = N^*{}_{\dot{\alpha}_1}{}^{\dot{\beta}_1} \ldots N^*{}_{\dot{\alpha}_m}{}^{\dot{\beta}_m}\bar{\chi}_{\dot{\beta}_1\ldots\dot{\beta}_m} \tag{5.65}$$

By taking the tensor product n tensor products of V acted on by the fundamental representations, with m tensor products of $\bar{V}$ acted on by the conjugate representation, one obtains a vector space which has elements with components $\psi_{\alpha_1\ldots\alpha_n\dot{\beta}_1\ldots\dot{\beta}_m}$ which transform as

$$\psi_{\alpha_1\ldots\alpha_n\dot{\beta}_1\ldots\dot{\beta}_m} \to \psi'_{\alpha_1\ldots\alpha_n\dot{\beta}_1\ldots\dot{\beta}_m} = N_{\alpha_1}{}^{\mu_1} \ldots N_{\alpha_n}{}^{\mu_n} N^*{}_{\dot{\beta}_1}{}^{\dot{\nu}_1} \ldots N^*{}_{\dot{\beta}_m}{}^{\dot{\nu}_m}\psi_{\beta_1\ldots\beta_n\dot{\nu}_1\ldots\dot{\nu}_m} \tag{5.66}$$

This representation is in general not irreducible.

5.4.1 Lorentz and $SL(2,\mathbb{C})$ indices

It is straightforward to map between Lorentz invariant tensors and $SL(2,\mathbb{C})$ invariant tensors. In particular, recall that the relationship between $N \in SL(2,\mathbb{C})$ and the corresponding Lorentz transformation $\Lambda = \Lambda(N)$ is given by

$$Nx_\mu\sigma^\mu N^\dagger = \eta_{\nu\rho}\Lambda^\rho{}_\gamma x^\gamma\sigma^\nu \tag{5.67}$$

which implies that

$$N\sigma^\mu N^\dagger = \eta_{\nu\rho}\Lambda^\rho{}_\gamma\eta^{\gamma\mu}\sigma^\nu = (\eta^{-1}\Lambda^T\eta)^\mu{}_\nu\sigma^\nu = (\Lambda^{-1})^\mu{}_\nu\sigma^\nu \tag{5.68}$$

So, denoting the components of σ^μ by $\sigma^\mu_{\alpha\dot\beta}$, one finds

$$\sigma^\nu_{\alpha\dot\beta} = N_\alpha{}^\lambda N^{*\dot\gamma}_{\dot\beta}\Lambda^\nu{}_\mu \sigma^\mu_{\lambda\dot\gamma} \tag{5.69}$$

which implies that $\sigma^\mu_{\alpha\dot\beta}$ is invariant. One can also define

$$(\bar\sigma^\mu)^{\dot\alpha\alpha} = \epsilon^{\alpha\beta}\epsilon^{\dot\alpha\dot\beta}\sigma^\mu_{\beta\dot\beta} \tag{5.70}$$

so that $\bar\sigma^0 = \sigma^0$, $\bar\sigma^i = -\sigma^i$ for $i = 1, 2, 3$.

Exercise: Prove the following useful identities

i) $\sigma^\mu\bar\sigma^\nu + \sigma^\nu\bar\sigma^\mu = 2\eta^{\mu\nu}\mathbb{I}_2$

ii) $\mathrm{Tr}\ \sigma^\mu\bar\sigma^\nu = 2\eta^{\mu\nu}$

iii) $\sigma^\mu_{\alpha\dot\alpha}\bar\sigma^{\dot\beta\beta}_\mu = 2\delta^\beta_\alpha\delta^{\dot\beta}_{\dot\alpha}$.

Definition 46. *Define the* 4×4 *matrices* γ^μ *by*

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar\sigma^\mu & 0 \end{pmatrix} \tag{5.71}$$

Then these matrices satisfy

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{I}_4 \tag{5.72}$$

An algebra satisfying this property is called a Clifford algebra.

Definition 47. *A Dirac spinor* Ψ_D *is a 4-component spinor constructed from left and right handed Weyl spinors* ψ_α, $\bar\chi^{\dot\alpha}$ *via*

$$\Psi_D = \begin{pmatrix} \psi_\alpha \\ \bar\chi^{\dot\alpha} \end{pmatrix} \tag{5.73}$$

The gamma matrices act on Dirac spinors.

The σ-matrix identities are useful. For example; recall that the correspondence between $\Lambda \in SO(3,1)^\uparrow$ and $N \in SL(2,\mathbb{C})$ is given by $N\sigma^\mu N^\dagger = \eta_{\nu\rho}\Lambda^\rho{}_\gamma\eta^{\gamma\mu}\sigma^\nu$. Then using (ii) above the components of Λ are given by

$$\Lambda^\mu{}_\nu = \frac{1}{2}\mathrm{Tr}\left(\bar\sigma^\mu N\sigma_\nu N^\dagger\right) \tag{5.74}$$

Also, it is straightforward to relate tensors with $SL(2,\mathbb{C})$ indices to tensors with Lorentz indices. Given a 4-vector with Lorentz indices V^μ one can define a tensor with $SL(2,\mathbb{C})$ indices via

$$V_{\alpha\dot\alpha} = V^\mu(\sigma_\mu)_{\alpha\dot\alpha} \tag{5.75}$$

The invariance of $(\sigma^\mu)_{\alpha\dot\alpha}$ ensures that if V^μ transforms as $V^\mu \to \Lambda^\mu{}_\nu V^\nu$ under the action of the Lorentz group, then $V_{\alpha\dot\alpha}$ transforms as $V_{\alpha\dot\alpha} \to N_\alpha{}^\beta N^{*\dot\beta}_{\dot\alpha} V_{\beta\dot\beta}$ under $SL(2,\mathbb{C})$. This expression can be inverted using (ii) of the above exercise to give

$$V^\mu = \frac{1}{2}(\bar\sigma^\mu)^{\dot\alpha\alpha}V_{\alpha\dot\alpha} \tag{5.76}$$

Similar maps between higher order tensors can also be constructed.

5.4.2 The Lie algebra of $SL(2,\mathbb{C})$

The Lie algebra of $SL(2,\mathbb{C})$ consists of traceless complex 2×2 matrices; which has six real dimensions. This is to be expected, as $\mathcal{L}(SO(3,1))$ is six-dimensional. It is convenient to define the matrices

$$\begin{aligned}
(\sigma^{\mu\nu})_\alpha{}^\beta &= \frac{i}{4}(\sigma^\mu\bar{\sigma}^\nu - \sigma^\nu\bar{\sigma}^\mu)_\alpha{}^\beta \\
(\bar{\sigma}^{\mu\nu})^{\dot{\alpha}}{}_{\dot{\beta}} &= \frac{i}{4}(\bar{\sigma}^\mu\sigma^\nu - \bar{\sigma}^\nu\sigma^\mu)^{\dot{\alpha}}{}_{\dot{\beta}}
\end{aligned} \tag{5.77}$$

so that $\sigma^{0i} = -\frac{i}{2}\sigma^i$, $\sigma^{jk} = \frac{1}{2}\epsilon^{jk\ell}\sigma^\ell$, $\bar{\sigma}^{0i} = \frac{i}{2}\sigma^i$, $\bar{\sigma}^{jk} = \frac{1}{2}\epsilon^{jk\ell}\sigma^\ell$. It is clear that the $\sigma^{\mu\nu}$ span the 2×2 traceless matrices over $\mathbb{R}$ (as do the $\bar{\sigma}^{\mu\nu}$), hence they are generators of the Lie algebra of $SL(2,\mathbb{C})$. By a direct computation we obtain the commutation relations:

$$[\sigma^{\mu\nu},\sigma^{\rho\sigma}] = i\big(\eta^{\nu\rho}\sigma^{\mu\sigma} + \eta^{\mu\sigma}\sigma^{\nu\rho} - \eta^{\nu\sigma}\sigma^{\mu\rho} - \eta^{\mu\rho}\sigma^{\nu\sigma}\big) \tag{5.78}$$

which is the same commutation relation as for the Lie algebra $\mathcal{L}(SO(3,1))$. Similarly, we find

$$[\bar{\sigma}^{\mu\nu},\bar{\sigma}^{\rho\sigma}] = i\big(\eta^{\nu\rho}\bar{\sigma}^{\mu\sigma} + \eta^{\mu\sigma}\bar{\sigma}^{\nu\rho} - \eta^{\nu\sigma}\bar{\sigma}^{\mu\rho} - \eta^{\mu\rho}\bar{\sigma}^{\nu\sigma}\big) \tag{5.79}$$

Hence the $\sigma^{\mu\nu}$ and $\bar{\sigma}^{\mu\nu}$ correspond to representations of $\mathcal{L}(SO(3,1))$.

The action of $SL(2,\mathbb{C})$ on left-handed Weyl spinors is given by

$$\psi_\alpha \to \big(e^{\omega_{\mu\nu}\sigma^{\mu\nu}}\big)_\alpha{}^\beta\psi_\beta \tag{5.80}$$

Just as for the Lorentz algebra, one can define $J_i = \frac{1}{2}\epsilon^{ijk}\sigma_{jk} = \frac{1}{2}\sigma^i$ and $K_i = \sigma_{0i} = \frac{i}{2}\sigma^i$. Hence $A_i = \frac{1}{2}\sigma^i$, $B_i = 0$. Therefore the fundamental representation corresponds to a spin-$\frac{1}{2}$ $\mathcal{L}(SU(2))$ representation generated by A, and a $\mathcal{L}(SU(2))$ B-singlet. This representation is denoted by $(\frac{1}{2},0)$.

The action of $SL(2,\mathbb{C})$ on right-handed Weyl spinors is given by

$$\bar{\chi}_{\dot{\alpha}} \to \big(e^{\omega_{\mu\nu}\bar{\sigma}^{\mu\nu}}\big)^{\dot{\beta}}{}_{\dot{\alpha}}\bar{\chi}_{\dot{\beta}} \tag{5.81}$$

Again, define $J_i = \frac{1}{2}\epsilon^{ijk}\bar{\sigma}_{jk} = \frac{1}{2}\sigma^i$ and $K_i = \bar{\sigma}_{0i} = -\frac{i}{2}\sigma^i$. Hence $A_i = 0$, $B_i = \frac{1}{2}\sigma^i$. Therefore this representation corresponds to a spin-$\frac{1}{2}$ $\mathcal{L}(SU(2))$ representation generated by B, and a $\mathcal{L}(SU(2))$ A-singlet. This representation is denoted by $(0,\frac{1}{2})$.

5.5 The Poincaré Group

The Poincaré group consists of Lorentz transformations combined with translations; which act on the spacetime co-ordinates by

$$x^\mu \to \Lambda^\mu{}_\nu x^\nu + b^\mu \tag{5.82}$$

where Λ is a Lorentz transformation and $b \in \mathbb{R}^4$ is an arbitrary 4-vector. One can denote the generic Poincaré group element by a pair (Λ, b) which act in this way. Note that under the action of (Λ, b) followed by (Λ', b'); $x \to \Lambda'\Lambda x + \Lambda' b + b'$, hence one defines the group product to be

$$(\Lambda', b')(\Lambda, b) = (\Lambda'\Lambda, \Lambda' b + b') \tag{5.83}$$

so the Poincareé group is closed under this multiplication. The identity is $(\mathbb{I}_4, 0)$ and the inverse of (Λ, b) is $(\Lambda^{-1}, -\Lambda^{-1}b)$.

One can construct a group isomorphism between the Poincaré group and the subgroup of $GL(5, \mathbb{R})$ of matrices of the form

$$\begin{pmatrix} \Lambda & b \\ 0 & 1 \end{pmatrix} \tag{5.84}$$

where Λ is a Lorentz transformation and b is an arbitrary 4-vector, as under matrix multiplication

$$\begin{pmatrix} \Lambda' & b' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \Lambda'\Lambda & \Lambda' b + b' \\ \underline{0} & 1 \end{pmatrix} \tag{5.85}$$

5.5.1 The Poincaré Algebra

Consider a curve

$$\begin{pmatrix} \Lambda(t) & b(t) \\ 0 & 1 \end{pmatrix} \tag{5.86}$$

in the Poincaré group passing through the identity when $t = 0$, so $\Lambda(0) = \mathbb{I}_4$, $b(0) = 0$. Differentiating with respect to t and setting $t = 0$ we note that the generic element of the Poincaré Lie algebra is of the form

$$\begin{pmatrix} m & v \\ 0 & 0 \end{pmatrix} \tag{5.87}$$

where $m \in \mathcal{L}(SO(3,1))$ and $v \in \mathbb{R}^4$ is unconstrained. Hence a basis for the Lie algebra is given by the 5×5 matrices $M^{\mu\nu}$ and P^ν for $\mu, \nu = 0, 1, 2, 3$ where

$$\begin{aligned} (M^{\rho\sigma})^\mu{}_\nu &= i(\eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho{}_\nu) \\ (M^{\rho\sigma})^4{}_\nu &= (M^{\rho\sigma})^\mu{}_4 = (M^{\rho\sigma})^4{}_4 = 0 \end{aligned} \tag{5.88}$$

and

$$\begin{aligned} (P^\nu)^\mu{}_4 &= i\eta^{\mu\nu} \\ (P^\nu)^\mu{}_\lambda &= (P^\nu)^4{}_\lambda = (P^\nu)^4{}_4 = 0 \end{aligned} \tag{5.89}$$

(labeling the matrix indices by $\mu, \nu = 0, 1, 2, 3$ and the additional index is "4"). The $M^{\rho\sigma}$ generate the Lorentz sub-algebra

$$\begin{pmatrix} m & 0 \\ 0 & 0 \end{pmatrix} \tag{5.90}$$

for $m \in \mathcal{L}(SO(3,1))$; they satisfy the usual Lorentz algebra commutation relations

$$[M^{\mu\nu}, M^{\rho\sigma}] = i\big(M^{\mu\sigma}\eta^{\nu\rho} + M^{\nu\rho}\eta^{\mu\sigma} - M^{\mu\rho}\eta^{\nu\sigma} - M^{\nu\sigma}\eta^{\mu\rho}\big) \tag{5.91}$$

The P^{ν} generate the translations

$$\begin{pmatrix} 0 & iv \\ 0 & 0 \end{pmatrix} \tag{5.92}$$

for $v \in \mathbb{R}^4$. The P^{μ} satisfy

$$[P^{\mu}, P^{\nu}] = 0 \tag{5.93}$$

and

$$[P^{\mu}, M^{\rho\sigma}] = i\eta^{\rho\mu}P^{\sigma} - i\eta^{\sigma\mu}P^{\rho} \tag{5.94}$$

The commutation relations

$$\begin{aligned} [M^{\mu\nu}, M^{\rho\sigma}] &= i\big(M^{\mu\sigma}\eta^{\nu\rho} + M^{\nu\rho}\eta^{\mu\sigma} - M^{\mu\rho}\eta^{\nu\sigma} - M^{\nu\sigma}\eta^{\mu\rho}\big) \\ [P^{\mu}, M^{\rho\sigma}] &= i\eta^{\rho\mu}P^{\sigma} - i\eta^{\sigma\mu}P^{\rho} \\ [P^{\mu}, P^{\nu}] &= 0 \end{aligned} \tag{5.95}$$

define the Poincaré algebra.

5.5.2 Representations of the Poincaré Algebra

Definition 48. *Suppose that d is a representation of the Poincaré algebra. Let $\epsilon_{\mu\nu\rho\sigma}$ be the totally antisymmetric tensor with $\epsilon_{0123} = 1$. Then the Pauli-Lubanski vector is defined by*

$$W_{\mu} = \frac{1}{2}\epsilon_{\mu\rho\sigma\nu} d(M^{\rho\sigma}) d(P^{\nu}) \tag{5.96}$$

Proposition 31. *The Pauli-Lubanski vector satisfies the following commutation relations:*

1) $[W_{\mu}, d(P_{\nu})] = 0$

2) $[W_{\mu}, d(M_{\rho\sigma})] = i\eta_{\mu\rho}W_{\sigma} - i\eta_{\mu\sigma}W_{\rho}$

3) $[W_{\mu}, W_{\nu}] = -i\epsilon_{\mu\nu\rho\sigma}W^{\rho}d(P^{\sigma})$

Proof

We will use the identities

$$\epsilon_{\mu\alpha\beta\gamma}\epsilon^{\mu\rho\sigma\tau} = -6\delta^{\rho}{}_{[\alpha}\delta^{\sigma}{}_{\beta}\delta^{\tau}{}_{\gamma]} = -6\delta^{[\rho}{}_{\alpha}\delta^{\sigma}{}_{\beta}\delta^{\tau]}{}_{\gamma} \tag{5.97}$$

and

$$\epsilon_{\mu\nu\alpha\beta}\epsilon^{\mu\nu\rho\sigma} = -4\delta^{\rho}{}_{[\alpha}\delta^{\sigma}{}_{\beta]} = -4\delta^{[\rho}{}_{\alpha}\delta^{\sigma]}{}_{\beta} \tag{5.98}$$

To prove (1) is straightforward:

$$[W_{\mu}, d(P_{\nu})] = \frac{1}{2}\epsilon_{\mu\rho\sigma\theta}[d(M^{\rho\sigma})d(P^{\theta}), d(P_{\nu})]$$

$$
\begin{aligned}
&= \frac{1}{2}\epsilon_{\mu\rho\sigma\theta}\big(d(M^{\rho\sigma})[d(P^\theta), d(P_\nu)] + [d(M^{\rho\sigma}), d(P_\nu)]d(P^\theta)\big) \\
&= \frac{1}{2}\epsilon_{\mu\rho\sigma\theta}\big(d(M^{\rho\sigma})d([P^\theta, P_\nu]) + d([M^{\rho\sigma}, P_\nu])d(P^\theta)\big) \\
&= \frac{1}{2}\epsilon_{\mu\rho\sigma\theta}\big(-i\delta^\rho{}_\nu d(P^\sigma) + i\delta^\sigma{}_\nu d(P^\rho)\big)d(P^\theta) \\
&= 0
\end{aligned}
\tag{5.99}
$$

To prove (ii) is an unpleasant exercise in algebra:

$$
\begin{aligned}
[W_\mu, d(M_{\rho\sigma})] &= \frac{1}{2}\epsilon_{\mu\lambda\chi\theta}[d(M^{\lambda\chi})d(P^\theta), d(M_{\rho\sigma})] \\
&= \frac{1}{2}\epsilon_{\mu\lambda\chi\theta}\big(d(M^{\lambda\chi})[d(P^\theta), d(M_{\rho\sigma})] + [d(M^{\lambda\chi}), d(M_{\rho\sigma})]d(P^\theta)\big) \\
&= \frac{1}{2}\epsilon_{\mu\lambda\chi\theta}\big(d(M^{\lambda\chi})d([P^\theta, M_{\rho\sigma}]) + d([M^{\lambda\chi}, M_{\rho\sigma}])d(P^\theta)\big) \\
&= \frac{1}{2}\epsilon_{\mu\lambda\chi\theta}\big(d(M^{\lambda\chi})\big(i\delta^\theta{}_\rho d(P_\sigma) - i\delta^\theta{}_\sigma d(P_\rho)\big) \\
&+ i\big(d(M^\lambda{}_\sigma)\delta^\chi{}_\rho - d(M^\chi{}_\sigma)\delta^\lambda{}_\rho - d(M^\lambda{}_\rho)\delta^\chi{}_\sigma + d(M^\chi{}_\rho)\delta^\lambda{}_\sigma\big)d(P^\theta)\big) \\
&= \frac{i}{2}\epsilon_{\mu\lambda\chi\theta}\big(d(M^{\lambda\chi})\delta^\theta{}_\rho d(P_\sigma) - d(M^{\lambda\chi})\delta^\theta{}_\sigma d(P_\rho) \\
&+ 2d(M^\lambda{}_\sigma)\delta^\chi{}_\rho d(P^\theta) - 2d(M^\lambda{}_\rho)\delta^\chi{}_\sigma d(P^\theta)\big) \\
&= \frac{i}{2}(\eta_{\sigma\tau}\delta^\theta{}_\rho - \eta_{\rho\tau}\delta^\theta{}_\sigma)\epsilon_{\mu\lambda\chi\theta}\big(d(M^{\lambda\chi})d(P^\tau) - 2d(M^{\lambda\tau})d(P^\chi)\big) \\
&= \frac{3i}{2}(\eta_{\sigma\tau}\delta^\theta{}_\rho - \eta_{\rho\tau}\delta^\theta{}_\sigma)\epsilon_{\mu\lambda\chi\theta}d(M^{[\lambda\chi})d(P^{\tau]})
\end{aligned}
\tag{5.100}
$$

But

$$
\begin{aligned}
\epsilon^{\lambda\chi\tau}{}_\gamma W^\gamma &= \frac{1}{2}\epsilon^{\lambda\chi\tau}{}_\gamma\epsilon^\gamma{}_{\nu_1\nu_2\nu_3}d(M^{\nu_1\nu_2})d(P^{\nu_3}) \\
&= 3d(M^{[\lambda\chi})d(P^{\tau]})
\end{aligned}
\tag{5.101}
$$

Hence

$$
\begin{aligned}
[W_\mu, d(M_{\rho\sigma})] &= \frac{i}{2}(\eta_{\sigma\tau}\delta^\theta{}_\rho - \eta_{\rho\tau}\delta^\theta{}_\sigma)\epsilon_{\mu\lambda\chi\theta}\epsilon^{\lambda\chi\tau}{}_\gamma W^\gamma \\
&= \frac{i}{2}(\eta_{\sigma\tau}\delta^\theta{}_\rho - \eta_{\rho\tau}\delta^\theta{}_\sigma)(-2)(\delta^\tau{}_\mu\eta_{\theta\gamma} - \delta^\tau{}_\theta\eta_{\mu\gamma})W^\gamma \\
&= i\eta_{\mu\rho}W_\sigma - i\eta_{\mu\sigma}W_\rho
\end{aligned}
\tag{5.102}
$$

as required.

(3) follows straightforwardly from (2):

$$
\begin{aligned}
[W_\mu, W_\nu] &= \frac{1}{2}\epsilon_{\nu\rho\sigma\theta}[W_\mu, d(M^{\rho\sigma})d(P^\theta)] \\
&= \frac{1}{2}\epsilon_{\nu\rho\sigma\theta}\big([W_\mu, d(M^{\rho\sigma})]d(P^\theta) + d(M^{\rho\sigma})[W_\mu, d(P^\theta)]\big) \\
&= \frac{1}{2}\epsilon_{\nu\rho\sigma\theta}[W_\mu, d(M^{\rho\sigma})]d(P^\theta)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2}\epsilon_{\nu\rho\sigma\theta}(i\delta^{\rho}{}_{\mu}W^{\sigma} - i\delta^{\sigma}{}_{\mu}W^{\rho})d(P^{\theta}) \\
&= -i\epsilon_{\mu\nu\rho\sigma}W^{\rho}d(P^{\sigma})
\end{aligned}
\tag{5.103}
$$

as required. ■.

From this we find the

Corollory 3. *The following commutation relations hold*

1) $[W_{\mu}W^{\mu}, d(P_{\nu})] = 0$

2) $[W_{\mu}W^{\mu}, d(M_{\rho\sigma})] = 0$

Proof

(1) follows because $[W_{\mu}W^{\mu}, d(P_{\nu})] = W_{\mu}[W^{\mu}, d(P_{\nu})] + [W_{\mu}, d(P_{\nu})]W^{\mu} = 0$

(2) holds because

$$
\begin{aligned}
[W_{\mu}W^{\mu}, d(M_{\rho\sigma})] &= W_{\mu}[W^{\mu}, d(M_{\rho\sigma})] + [W_{\mu}, d(M_{\rho\sigma})]W^{\mu} \\
&= W_{\mu}(i\delta^{\mu}{}_{\rho}W_{\sigma} - i\delta^{\mu}{}_{\sigma}W_{\rho}) + (i\eta_{\mu\rho}W_{\sigma} - i\eta_{\mu\sigma}W_{\rho})W^{\mu} \\
&= 0
\end{aligned}
\tag{5.104}
$$

as required. ■

Hence we have shown that $W_{\mu}W^{\mu}$ is a Casimir operator. $d(P_{\mu})d(P^{\mu})$ is another Casimir operator:

Proposition 32. *The following commutation relations hold*

1) $[d(P_{\mu})d(P^{\mu}), d(P_{\nu})] = 0$

2) $[d(P_{\mu})d(P^{\mu}), d(M_{\rho\sigma})] = 0$

Proof

(1) follows because

$$
[d(P_{\mu})d(P^{\mu}), d(P_{\nu})] = d(P_{\mu})[d(P^{\mu}), d(P_{\nu})] + [d(P_{\mu}), d(P_{\nu})]d(P^{\mu}) = 0 \tag{5.105}
$$

(2) holds because

$$
\begin{aligned}
[d(P_{\mu})d(P^{\mu}), d(M_{\rho\sigma})] &= d(P_{\mu})[d(P^{\mu}), d(M_{\rho\sigma})] + [d(P_{\mu}), d(M_{\rho\sigma})]d(P^{\mu}) \\
&= d(P_{\mu})(i\delta^{\mu}{}_{\rho}d(P_{\sigma}) - i\delta^{\mu}{}_{\sigma}d(P_{\rho})) + (i\eta_{\mu\rho}d(P_{\sigma}) - i\eta_{\mu\sigma}d(P_{\rho}))d(P^{\mu}) \\
&= 0
\end{aligned}
\tag{5.106}
$$

as required. ■

We shall show that irreducible representations are classified by the values of the two Casimir operators $W_{\mu}W^{\mu}$ and $d(P_{\mu})d(P^{\mu})$.

In particular, suppose that $\mathcal{D}$ is a unitary representation of the Poincaré group acting on V. Such representations arise naturally in the context of quantum field theory when V is taken to be a Hilbert space, and it is assumed that Poincaré transformations do not affect transition probabilities. We will assume that this is the case.

Note that $iM_{\mu\nu}$ and iP^μ form a basis for the (real) Poincaré algebra. Hence one can locally write the Poincaré transformation as

$$e^{-\frac{i}{2}(b_\mu P^\mu + \omega_{\mu\nu} M^{\mu\nu})} \tag{5.107}$$

for real b_μ and skew-symmetric real $\omega_{\mu\nu}$, and

$$\mathcal{D}(e^{-\frac{i}{2}(b_\mu P^\mu + \omega_{\mu\nu} M^{\mu\nu})}) = e^{-\frac{i}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))} \tag{5.108}$$

where d is a representation of the Poincaré algebra acting on V. As $\mathcal{D}$ is unitary, $d(M_{\rho\sigma})$ and $d(P^\mu)$ are hermitian.

As the $d(P^\mu)$ commute with each other and are hermitian, they can be simultaneously diagonalized, with real eigenvalues. For a 4-vector q^μ define the subspace V_q of V to be the simultaneous eigenspace

$$V_q = \{|\psi\rangle \in V : d(P^\mu)\,|\psi\rangle = q^\mu\,|\psi\rangle\,, \quad \mu = 0,1,2,3\} \tag{5.109}$$

and

$$V = \bigoplus_q V_q \tag{5.110}$$

Then on V_q, $d(P_\mu)d(P^\mu) = q^\mu q_\mu = q^2$. We will assume that for configurations of physical interest, such as when q is the 4-momentum of a massive particle or of a photon, that $q^2 \geq 0$ and $q^0 > 0$. We will only consider these cases.

Consider first the operators

$$h^\lambda(t) = e^{\frac{it}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))} d(P^\lambda) e^{-\frac{it}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))} \tag{5.111}$$

Differentiating with respect to t we find

$$\begin{aligned}\frac{dh^\lambda}{dt} &= e^{\frac{it}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))}[d(P^\lambda), -\frac{i}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))]e^{-\frac{it}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))} \\ &= \omega^\lambda{}_\chi e^{\frac{it}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))} d(P^\chi) e^{-\frac{it}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))} \\ &= \omega^\lambda{}_\chi h^\chi \end{aligned} \tag{5.112}$$

with the initial condition $h^\lambda(0) = d(P^\lambda)$. Therefore

$$h^\lambda(t) = (e^{t\omega})^\lambda{}_\rho d(P^\rho) = (e^{-\frac{it}{2}\omega_{\mu\nu} M^{\mu\nu}})^\lambda{}_\rho d(P^\rho) \tag{5.113}$$

Hence, setting $t = 1$ we find

$$d(P^\lambda) e^{-\frac{i}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))} = (e^{-\frac{i}{2}\omega_{\mu\nu} M^{\mu\nu}})^\lambda{}_\rho e^{-\frac{i}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))} d(P^\rho) \tag{5.114}$$

So, if $|\psi\rangle \in V_q$ then $e^{-\frac{i}{2}(b_\mu d(P^\mu) + \omega_{\mu\nu} d(M^{\mu\nu}))}\,|\psi\rangle \in V_{q'}$ where $q' = e^{-\frac{i}{2}\omega_{\mu\nu} M^{\mu\nu}} q$.

Definition 49. *The stability subgroup H_q (or "little group") which is associated with V_q is the subgroup of the Poincaré group defined by*

$$H_q = \{e^{-\frac{i}{2}(b_\mu P^\mu + \omega_{\mu\nu}M^{\mu\nu})} : e^{-\frac{i}{2}(b_\mu d(P^\mu)+\omega_{\mu\nu}d(M^{\mu\nu}))}\,|\psi\rangle \in V_q \quad \text{for } |\psi\rangle \in V_q\} \tag{5.115}$$

It can be shown that H_q is a Lie subgroup of the Poincaré group. Suppose then that $-\frac{i}{2}(b_\mu P^\mu + \omega_{\mu\nu}M^{\mu\nu}) \in \mathcal{L}(H_q)$. It follows that $e^{-\frac{it}{2}\omega_{\mu\nu}M^{\mu\nu}}q = q$ for $t \in \mathbb{R}$. Expanding out in powers of t we see that this constraint is equivalent to

$$\omega_{\mu\nu}q^\nu = 0 \tag{5.116}$$

which has a general solution

$$\omega_{\mu\nu} = \epsilon_{\mu\nu\rho\sigma}n^\rho q^\sigma \tag{5.117}$$

where n^ρ is an arbitrary constant 4-vector. Hence, if $|\psi\rangle \in V_q$ and $e^{-\frac{i}{2}(b_\mu P^\mu+\omega_{\mu\nu}M^{\mu\nu})} \in H_q$ then

$$e^{-\frac{i}{2}(b_\mu d(P^\mu)+\omega_{\mu\nu}d(M^{\mu\nu}))}\,|\psi\rangle = e^{-\frac{i}{2}b_\mu q^\mu}e^{-in^\mu W_\mu}\,|\psi\rangle \tag{5.118}$$

so we have reduced the action of H_q on V_q to the action of W_μ on V_q.

The action of a *generic* Poincaré transformation (Λ, b) does not leave V_q invariant, because

$$V_q \to V'_q = V_{\Lambda q} \tag{5.119}$$

as $q' = \Lambda q$. However, $q^2 = q'^2$ is invariant. Hence we can split V into invariant subspaces V_q corresponding to vectors q^μ *which have the same value of* q^2. We will therefore henceforth work with such an invariant subspace, and consider $q^2 = m^2$ to be fixed.

If $m^2 > 0$ then there is a Lorentz transformation $\Lambda'(q)$ such that $q = \Lambda'(q)k$ where $k^\mu = (m, 0, 0, 0)$. Alternatively, if $m = 0$ then there is a Lorentz transformation $\Lambda(q)$ such that $q = \Lambda'(q)k$ where $k^\mu = (E, E, 0, 0)$. These Lorentz transformations can be taken to be fixed functions of the q.

The key step is to show that the action of the entire Poincaré group on the V_q (with $q^2 = m^2$ fixed) is fixed by the action of H_k on V_k (which is in turn determined by the action of the Pauli-Lubanski vector on V_k).

To show this, first note that if $|\psi_k\rangle \in V_k$, then one can write

$$|\psi_q\rangle = \mathcal{D}(\Lambda'(q), 0)\,|\psi_k\rangle \tag{5.120}$$

where $|\psi_q\rangle \in V_q$, and these transformations can be used to obtain all elements of V_q from those in V_k.

It is then straightforward to show that the action of the representation of the whole Poincaré group on $\{V_q : q^2 = m^2\}$ is determined by the action of $\mathcal{D}(H_k)$ acting on V_k.

To see this explicitly, suppose that $|\psi_{k,M}\rangle$ for $M = 1, \ldots, \ell_k$ is a basis of V_k. Then one can define

$$|\psi_{q,M}\rangle = \mathcal{D}(\Lambda'(q), 0)\,|\psi_{k,M}\rangle \tag{5.121}$$

and the $|\psi_{q,M}\rangle$ then form a basis for V_q. The representation of H_k on V_k is determined by the coefficients $\mathcal{D}(h)_{MN}$, where $h \in H_k$, in the expansion

$$\mathcal{D}(h)\,|\psi_{k,M}\rangle = \sum_N \mathcal{D}(h)_{MN}\,|\psi_{k,N}\rangle \tag{5.122}$$

Suppose that (Λ, b) is a generic Poincaré transformation; then one can write

$$\begin{aligned}\mathcal{D}(\Lambda,b)\,|\psi_{q,M}\rangle &= \mathcal{D}(\Lambda,b)\mathcal{D}(\Lambda'(q),0)\,|\psi_{k,M}\rangle \\ &= \mathcal{D}(\Lambda'(\Lambda q),0)\mathcal{D}(\Lambda'(\Lambda q)^{-1}\Lambda\Lambda'(q),\Lambda'(\Lambda q)^{-1}b)\,|\psi_{k,M}\rangle\end{aligned} \tag{5.123}$$

However, $((\Lambda'(\Lambda q))^{-1}\Lambda\Lambda'(q), \Lambda'(\Lambda q)^{-1}b) \in H_k$, so it is possible to expand

$$\begin{aligned}\mathcal{D}(\Lambda,b)\,|\psi_{q,M}\rangle &= \sum_N \mathcal{D}(\Lambda'(\Lambda q),0)\mathcal{D}(\Lambda'(\Lambda q)^{-1}\Lambda\Lambda'(q),\Lambda'(\Lambda q)^{-1}b)_{MN}\,|\psi_{k,N}\rangle \\ &= \sum_N \mathcal{D}(\Lambda'(\Lambda q)^{-1}\Lambda\Lambda'(q),\Lambda'(\Lambda q)^{-1}b)_{MN}\,|\psi_{\Lambda q,N}\rangle\end{aligned} \tag{5.124}$$

We will examine the action of H_k on V_k in the timelike and null cases separately. Although

$$V = \bigoplus_q V_q \tag{5.125}$$

is not in general finite-dimensional, we shall assume that V_k (and hence the V_q) are finite dimensional.

5.5.3 Massive Representations of the Poincaré Group: $k^\mu = (m,0,0,0)$

We compute the action of W_μ on V_k. If $|\psi\rangle \in V_k$ then

$$\begin{aligned}W_0\,|\psi\rangle &= \frac{1}{2}\epsilon_{ij\ell}d(M^{ij})d(P^\ell)\,|\psi\rangle = 0 \\ W_i\,|\psi\rangle &= -\frac{1}{2}\epsilon_{ij\ell}d(M^{j\ell})d(P^0)\,|\psi\rangle = -md(J_i)\,|\psi\rangle\end{aligned} \tag{5.126}$$

We have already shown that $d(J_i)$ generates a $\mathcal{L}(SU(2))$ algebra, hence the little group for massive representations is $SO(3)$. For irreducible representations, the spin is fixed by the value taken by the Casimir on V_k; $W_\mu W^\mu\,|\psi\rangle = -m^2 d(J_i)d(J_i)\,|\psi\rangle$.

5.5.4 Massless Representations of the Poincaré Group: $k^\mu = (E,E,0,0)$

Again, we compute the action of W_μ on V_k. If $|\psi\rangle \in V_k$ then

$$\begin{aligned}W_0\,|\psi\rangle &= \frac{1}{2}\epsilon_{ij\ell}d(M^{ij})d(P^\ell)\,|\psi\rangle = d(M^{23})d(P^1)\,|\psi\rangle = Ed(J_1)\,|\psi\rangle \\ W_1\,|\psi\rangle &= -Ed(M^{23})\,|\psi\rangle = -Ed(J_1)\,|\psi\rangle \\ W_2\,|\psi\rangle &= \frac{1}{2}\epsilon_{2\mu\nu\lambda}d(M^{\mu\nu})d(P^\lambda)\,|\psi\rangle\end{aligned}$$

$$
\begin{aligned}
&= E(d(M^{13}) - d(M^{03}))\,|\psi\rangle \\
&= E(-d(J_2) + d(K_3))\,|\psi\rangle \\
W_3\,|\psi\rangle &= \frac{1}{2}\epsilon_{3\mu\nu\lambda} d(M^{\mu\nu}) d(P^\lambda)\,|\psi\rangle \\
&= E(-d(M^{12}) + d(M^{02}))\,|\psi\rangle \\
&= E(-d(J_3) - d(K_2))\,|\psi\rangle
\end{aligned} \tag{5.127}
$$

Observe that the following commutation relations hold:

$$
\begin{aligned}
[-d(J_2) + d(K_3), -d(J_3) - d(K_2)] &= 0 \\
[d(J_1), -d(J_3) - d(K_2)] &= -i(-d(J_2) + d(K_3)) \\
[d(J_1), -d(J_2) + d(K_3)] &= i(-d(J_3) - d(K_2))
\end{aligned} \tag{5.128}
$$

These expressions may be simplified slightly by setting $R_1 = -d(J_3) - d(K_2)$, $R_2 = -d(J_2) + d(K_3)$, $J = d(J_1)$; so that

$$
[R_1, R_2] = 0, \qquad [J, R_1] = -iR_2, \qquad [J, R_2] = iR_1 \tag{5.129}
$$

R_1 and R_2 are commuting hermitian operators on V_k, and hence can be simultaneously diagonalized over $\mathbb{R}$. Consider a state $|\psi\rangle \in V_k$ with $R_1\,|\psi\rangle = r_1\,|\psi\rangle$, $R_2\,|\psi\rangle = r_2\,|\psi\rangle$ for $r_1, r_2 \in \mathbb{R}$. Define

$$
\begin{aligned}
f(\theta) &= e^{-i\theta J} R_1 e^{i\theta J}\,|\psi\rangle \\
g(\theta) &= e^{-i\theta J} R_2 e^{i\theta J}\,|\psi\rangle
\end{aligned} \tag{5.130}
$$

for $\theta \in \mathbb{R}$. Differentiating with respect to θ and using the commutation relations we find $\frac{df}{d\theta} = -g$, $\frac{dg}{d\theta} = f$. Solving these equations with the initial condition $f(0) = r_1\,|\psi\rangle$, $g(0) = r_2\,|\psi\rangle$ we find

$$
\begin{aligned}
f(\theta) &= (r_1 \cos\theta - r_2 \sin\theta)\,|\psi\rangle \\
g(\theta) &= (r_1 \sin\theta + r_2 \cos\theta)\,|\psi\rangle
\end{aligned} \tag{5.131}
$$

which implies

$$
\begin{aligned}
R_1 e^{i\theta J}\,|\psi\rangle &= (r_1 \cos\theta - r_2 \sin\theta) e^{i\theta J}\,|\psi\rangle \\
R_2 e^{i\theta J}\,|\psi\rangle &= (r_1 \sin\theta + r_2 \cos\theta) e^{i\theta J}\,|\psi\rangle
\end{aligned} \tag{5.132}
$$

Hence, unless $r_1 = r_2 = 0$, there is a continuum of R_1, R_2 eigenstates which implies V_k cannot be finite-dimensional. We must therefore have $R_1 = R_2 = 0$ on V_k. J is also a hermitian operator on V_k, and can also be diagonalized. For irreducible representations, J can have only one eigenvalue, $\sigma \in \mathbb{R}$. σ is called the helicity of the particle. It follows that $W_\mu = \sigma k_\mu$, so it is clear that σ is a Lorentz invariant quantity.

There is no algebraic constraint fixing the value of the helicity σ in the massless case, as there is to fix the spin in the massive case. However, for physically realistic systems, one can make a topological argument to fix $2\sigma \in \mathbb{Z}$. This is because $e^{i\theta J}$ describes a rotation of angle θ in the spatial plane in the $2,3$ directions. So in particular, setting $\theta = 2\pi$ we find

$$
e^{2\pi i J}\,|\psi\rangle = e^{2\pi\sigma i}\,|\psi\rangle \tag{5.133}
$$

We require that $e^{2\pi\sigma i} = \pm 1$ (for a projective representation) and so $2\sigma \in \mathbb{Z}$. Neutrinos have helicity $\pm\frac{1}{2}$, photons have helicity ± 1 and gravitons have helicity ± 2.

6. Gauge Theories

Lie groups and Lie algebras play an important role in the various dynamical theories which govern the behaviour of particles - the gauge theories. Though we will not examine the quantization of these theories, we shall present the relationship between Lie algebras and gauge theories.

Before examining non-Abelian gauge theories, we briefly recap some properties of the simplest gauge theory, which is the $U(1)$ gauge theory of electromagnetism.

6.1 Electromagnetism

The gauge theory of electromagnetism contains a field strength

$$f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu \tag{6.1}$$

where $\mu, \nu = 0, 1, 2, 3$ and x^μ are co-ordinates on Minkowski space (indices raised/lowered with the Minkowski metric η), and $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$. $a_\mu \in \mathbb{R}$ is the 4-vector potential.

Under a gauge transformation $a_\mu \longrightarrow a'_\mu = a_\mu - \partial_\mu \lambda$ where λ is a real function,

$$f_{\mu\nu} \to f'_{\mu\nu} = \partial_\mu(a_\nu - \partial_\nu \lambda) - \partial_\nu(a_\mu - \partial_\mu \lambda) = \partial_\mu a_\nu - \partial_\nu a_\mu = f_{\mu\nu} \tag{6.2}$$

since $\partial_\mu \partial_\nu \lambda = \partial_\nu \partial_\mu \lambda$. Hence $f_{\mu\nu}$ is invariant under gauge transformations.

The field equations of electromagnetism are

$$\partial^\mu f_{\mu\nu} = j_\nu \tag{6.3}$$

and

$$\epsilon^{\lambda\mu\nu\rho} \partial_\mu f_{\nu\rho} = 0 \tag{6.4}$$

Equation (6.4) holds automatically due to the existence of the vector potential. Conversely, if $f_{\mu\nu}$ satisfies (6.4) then it can be shown that a vector potential a_μ exists (though only locally) such that $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$.

Using the vector potential, one defines a covariant derivative D_μ by

$$D_\mu \psi = \partial_\mu \psi + i a_\mu \psi \tag{6.5}$$

where $\psi = \psi(x)$. Under a gauge transformation

$$\psi \longrightarrow \psi' = e^{i\lambda}\psi, \qquad a_\mu \longrightarrow a'_\mu = a_\mu - \partial_\mu \lambda \tag{6.6}$$

where $\lambda = \lambda(x)$, it is straightforward to see that

$$\begin{aligned} D_\mu \psi \longrightarrow (D_\mu \psi)' &= \partial_\mu(e^{i\lambda}\psi) + i(a_\mu - \partial_\mu \lambda)e^{i\lambda}\psi \\ &= e^{i\lambda} D_\mu \psi \end{aligned} \tag{6.7}$$

so that $D_\mu \psi$ transforms like ψ. This means that the Dirac equation

$$i\gamma^\mu D_\mu \Psi - m\Psi = 0 \tag{6.8}$$

is gauge invariant. Ψ is a 4-component Dirac spinor constructed from left and right handed Weyl spinors ψ_α, $\bar{\chi}^{\dot{\alpha}}$ via

$$\Psi = \begin{pmatrix} \psi_\alpha \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix} \tag{6.9}$$

and the 4×4 matrices γ^μ are given by

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \tag{6.10}$$

These matrices satisfy the *Clifford algebra*

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{I}_4 \ . \tag{6.11}$$

The standard Lagrangian governing the interaction of electrodynamics with scalar fields

$$\mathcal{L} = -\frac{1}{4}f_{\mu\nu}f^{\mu\nu} + \frac{1}{2}(D_\mu\phi)^* D^\mu\phi + V(\phi^*\phi) \tag{6.12}$$

where ϕ is a complex scalar field, and V is a real function of $\phi^*\phi$ is also gauge covariant.

It is possible to obtain the gauge field strength in a natural way from the commutator of covariant derivatives. If ϕ is a scalar field then

$$\begin{aligned} D_\mu D_\mu\phi &= (\partial_\mu + ia_\mu)(\partial_\nu\phi + ia_\nu\phi) \\ &= \partial_\mu\partial_\nu\phi + i(a_\nu\partial_\mu\phi + a_\mu\partial_\nu\phi) - a_\mu a_\nu\phi + i\partial_\mu a_\nu\phi \end{aligned} \tag{6.13}$$

and hence

$$[D_\mu, D_\nu]\phi = i(\partial_\mu a_\nu - \partial_\nu a_\mu)\phi = if_{\mu\nu}\phi \tag{6.14}$$

6.2 Non-Abelian Gauge Theory

6.2.1 The Fundamental Covariant Derivative

Suppose that G is a compact matrix Lie group acting on a vector space V via the fundamental representation. Consider a scalar field $\Phi(x)$ which is an x^μ dependent element of V (which can be thought of as a column vector of ordinary scalar fields). Suppose that Φ transforms under the fundamental representation as

$$\Phi(x) \longrightarrow \Phi'(x) = g(x)\Phi(x) \tag{6.15}$$

where $g(x) \in G$.

Definition 50. *The fundamental covariant derivative D_μ is defined by*

$$D_\mu\Phi = \partial_\mu\Phi + A_\mu\Phi \tag{6.16}$$

where $A_\mu \in \mathcal{L}(G)$ is an element of the Lie algebra of G acting on V.

We require that $D_\mu\Phi$ should transform under local gauge transformations in the same way as Φ. Suppose that $A_\mu \to A'_\mu$ under local gauge transformations. Then we need

$$\partial_\mu\Phi' + A'_\mu\Phi' = g(\partial_\mu\Phi + A_\mu\Phi) \tag{6.17}$$

which implies

$$\partial_\mu(g\Phi) + A'_\mu g\Phi = g(\partial_\mu\Phi + A_\mu\Phi) \tag{6.18}$$

and hence

$$\partial_\mu g\Phi + A'_\mu g\phi = gA_\mu\Phi \tag{6.19}$$

As this must hold for all Φ, we find the transformation rule

$$A'_\mu = gA_\mu g^{-1} - \partial_\mu g g^{-1} \tag{6.20}$$

Before proceeding further, there is a question of consistency: namely if $A_\mu \in \mathcal{L}(G)$ then we must verify that A'_μ given above is also an element of $\mathcal{L}(G)$. This is proved using the

Lemma 11. *If $g(t)$ is a curve in the matrix Lie group G then $\frac{dg}{dt}g(t)^{-1} \in \mathcal{L}(G)$.*

Proof

Suppose that $g(t) = g_0$ when $t = t_0$. Set $h(t) = g(t+t_0)g_0^{-1}$. Then $h(t)$ is a smooth curve in G with $h(0) = \mathbb{I}$, and

$$\frac{dh}{dt}|_{t=0} = \frac{dg}{dt}|_{t=t_0}g_0^{-1} = \big(\frac{dg}{dt}g^{-1}\big)|_{t=t_0} \tag{6.21}$$

But $\frac{dh}{dt}|_{t=0} \in \mathcal{L}(G)$ by definition, and hence $\big(\frac{dg}{dt}g^{-1}\big)|_{t=t_0} \in \mathcal{L}(G)$ for all t_0. ■

Hence we have shown that $\partial_\mu g g^{-1} \in \mathcal{L}(G)$, and from our previous analysis of the adjoint representation, we know that $gA_\mu g^{-1} \in \mathcal{L}(G)$; so $A'_\mu \in \mathcal{L}(G)$ as required.

6.2.2 Generic Covariant Derivative

Definition 51. *Suppose that G is a matrix Lie group with representation $\mathcal{D}$ acting on V, and let d denote the associated representation of the Lie algebra acting on V. Let elements $\theta \in V$ transform as $\theta \to \theta' = \mathcal{D}(g(x))\theta$ under local gauge transformations.*

Then the covariant derivative D_μ associated with $\mathcal{D}$ acting on V is defined by

$$D_\mu\theta = \partial_\mu\theta + d(A_\mu)\theta \tag{6.22}$$

where $A_\mu \in \mathcal{L}(G)$ transforms as $A_\mu \to A'_\mu = gA_\mu g^{-1} - \partial_\mu g g^{-1}$.

In order to show that this covariant derivative transforms as $D_\mu\theta \to \mathcal{D}(g)D_\mu\theta$ we must prove the

Lemma 12. *Suppose that $\mathcal{D}$ is a representation of G acting on V, with associated representation d of $\mathcal{L}(G)$ acting on V. Then*

i) If $v \in \mathcal{L}(G)$ and $g \in G$, then $d(gvg^{-1})\mathcal{D}(g) = \mathcal{D}(g)d(v)$

ii) If $g(t)$ is a curve in G then $\frac{d\mathcal{D}(g)}{dt}\mathcal{D}(g^{-1}) = d\big(\frac{dg}{dt}g^{-1}\big)$

Proof

(*Caveat: in this proof t is simply a parameter along a curve in G, not the spacetime co-ordinate x^0!*)

To prove (i), set $g = e^h$ for $h \in \mathcal{L}(G)$, so that $\mathcal{D}(g) = e^{d(h)}$. Then (i) is equivalent to

$$e^{-d(h)} d(e^h v e^{-h}) e^{d(h)} = d(v) \tag{6.23}$$

Set

$$f(t) = e^{-td(h)} d(e^{th} v e^{-th}) e^{td(h)} \tag{6.24}$$

for $t \in \mathbb{R}$. Then

$$\begin{aligned} \frac{df}{dt} &= e^{-td(h)} \big([d(e^{th} v e^{-th}), d(h)] + d(e^{th}[h, v]e^{-th})\big) e^{td(h)} \\ &= e^{-td(h)} \big(d([e^{th} v e^{-th}, h]) + d(e^{th}[h, v]e^{-th})\big) e^{td(h)} \\ &= e^{-td(h)} \big(d(e^{th}[v, h]e^{-th}) + d(e^{th}[h, v]e^{-th})\big) e^{td(h)} \\ &= 0 \end{aligned} \tag{6.25}$$

and $f(0) = d(v)$. Hence, $f(1) = e^{-d(h)} d(e^h v e^{-h}) e^{d(h)} = f(0) = d(v)$ as required.

To prove (ii), suppose $g(t) = g_0$ at $t = t_0$. Set $h(t) = g(t + t_0) g_0^{-1}$, so that $h(t)$ is a smooth curve in G with $h(0) = \mathbb{I}$. Then

$$\frac{d\mathcal{D}(g)}{dt} \mathcal{D}(g^{-1})|_{t=t_0} = \frac{d\mathcal{D}(h(t)g_0)}{dt}|_{t=0} \mathcal{D}(g_0^{-1}) = \frac{d\mathcal{D}(h(t))}{dt}|_{t=0} \tag{6.26}$$

and

$$d(\frac{dg}{dt} g^{-1}|_{t=t_0}) = d(\frac{dh(t)}{dt}|_{t=0}) \tag{6.27}$$

As $h(0) = \mathbb{I}$ we can set $h(t) = e^{th_1 + O(t^2)}$ for some constant matrix h_1. Then

$$\frac{dh(t)}{dt}|_{t=0} = h_1 \in \mathcal{L}(G) \tag{6.28}$$

so

$$d(\frac{dh(t)}{dt}|_{t=0}) = d(h_1) \tag{6.29}$$

But by definition

$$d(h_1) = \frac{d}{dt}\big(\mathcal{D}(e^{th_1})\big)|_{t=0} = \frac{d}{dt}\big(\mathcal{D}(h(t))\big)|_{t=0} \tag{6.30}$$

Therefore

$$\frac{d\mathcal{D}(g)}{dt} \mathcal{D}(g^{-1})|_{t=t_0} = d(h_1) = d(\frac{dg}{dt} g^{-1}|_{t=t_0}) \tag{6.31}$$

are required. ■

Proposition 33. *The covariant derivative D_μ associated with $\mathcal{D}$ transforms as*

$$D_\mu \theta \to (D_\mu \theta)' = \mathcal{D}(g) D_\mu \theta \tag{6.32}$$

under local gauge transformations.

<u>**Proof**</u>

Note that

$$\partial_\mu \theta' + d(A'_\mu)\theta' = \partial_\mu(\mathcal{D}(g)\theta) + d(gA_\mu g^{-1} - \partial_\mu g g^{-1})\mathcal{D}(g)\theta$$

$$
\begin{aligned}
&= \mathcal{D}(g)(\partial_\mu\theta + d(A_\mu)\theta) + \big(d(gA_\mu g^{-1})\mathcal{D}(g) - \mathcal{D}(g)d(A_\mu) \\
&+ \partial_\mu(\mathcal{D}(g)) - d(\partial_\mu g g^{-1})\mathcal{D}(g)\big)\theta \\
&= \mathcal{D}(g)D_\mu\theta + \big(d(gA_\mu g^{-1})\mathcal{D}(g) - \mathcal{D}(g)d(A_\mu)\big)\theta \\
&+ \big(\partial_\mu(\mathcal{D}(g)) - d(\partial_\mu g g^{-1})\mathcal{D}(g)\big)\theta
\end{aligned} \tag{6.33}
$$

However, by the previous lemma, we have proved that $d(gA_\mu g^{-1})\mathcal{D}(g) - \mathcal{D}(g)d(A_\mu) = 0$ and $\partial_\mu(\mathcal{D}(g)) - d(\partial_\mu g g^{-1})\mathcal{D}(g) = 0$, hence

$$
\partial_\mu\theta' + d(A'_\mu)\theta' = \mathcal{D}(g)D_\mu\theta \tag{6.34}
$$

as required. ■

Given this property of transformations of covariant derivatives, one can define the adjoint covariant derivative

Definition 52. *Suppose that $\theta \in \mathcal{L}(G)$ transforms under the adjoint representation Ad of G. The the covariant derivative associated with the adjoint representation is*

$$
D_\mu\theta = \partial_\mu\theta + (\mathrm{ad}\ A_\mu)\theta = \partial_\mu\theta + [A_\mu, \theta] \tag{6.35}
$$

To summarize, we have shown that if Φ transforms under the action of the fundamental representation as $\Phi \to \Phi' = g\Phi$, then in order for the fundamental covariant derivative to transform in the same way, one must impose the transformation

$$
A_\mu \to A'_\mu = gA_\mu g^{-1} - \partial_\mu g g^{-1} \tag{6.36}
$$

on the gauge potential. We then have shown that if Φ transforms under the action of a *generic* representation $\mathcal{D}$, $\Phi \to \Phi' = \mathcal{D}(g)\Phi$, then the same transformation rule $A_\mu \to A'_\mu = gA_\mu g^{-1} - \partial_\mu g g^{-1}$ is sufficient to ensure that the *generic* covariant derivative $D_\mu\Phi = \partial_\mu\Phi + d(A_\mu)\Phi$ also transforms in the same way as Φ. **Caveat**: *a covariant derivative is always defined with respect to a particular representation*

6.3 Non-Abelian Yang-Mills Fields

Following from the relationship of the $U(1)$ electromagnetic field strength with the commutator of the $U(1)$ covariant derivatives acting on scalars, we consider the commutator of the fundamental covariant derivative D_μ acting on $\Phi(x) \in V$, which transforms under the fundamental representation as $\Phi \to \Phi' = g\Phi$:

$$
\begin{aligned}
[D_\mu, D_\nu]\Phi &= (\partial_\mu + A_\mu)\big(\partial_\nu\Phi + A_\nu\Phi\big) - (\partial_\nu + A_\nu)\big(\partial_\mu\Phi + A_\mu\Phi\big) \\
&= \big(\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]\big)\Phi
\end{aligned} \tag{6.37}
$$

Definition 53. *The non-abelian Yang-Mills field strength is*

$$
F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \tag{6.38}
$$

Note that as $A_\mu \in \mathcal{L}(g)$ it follows that $F_{\mu\nu} \in \mathcal{L}(G)$. Note also that by construction $[D_\mu, D_\nu]\Phi$ transforms like Φ under a gauge transformation. Hence if $F'_{\mu\nu}$ is the transformed gauge field strength, then $F'_{\mu\nu}\Phi' = gF_{\mu\nu}\Phi$. As this must hold for all Φ, we find

$$F'_{\mu\nu} = gF_{\mu\nu}g^{-1} \tag{6.39}$$

so that F transforms like the homogeneous part of A_μ.

Exercise: Verify this transformation rule for $F_{\mu\nu}$ directly from the definition $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$ together with the transformation rule of A_μ.

Lemma 13. *The non-abelian field strength satisfies*

$$D_\mu D_\nu F^{\mu\nu} = 0 \tag{6.40}$$

where here D_μ is the adjoint covariant derivative.

Proof

$$\begin{aligned} D_\mu D_\nu F^{\mu\nu} &= D_\mu(\partial_\nu F^{\mu\nu} + [A_\nu, F^{\mu\nu}]) \\ &= \partial_\mu\partial_\nu F^{\mu\nu} + [A_\mu, \partial_\nu F^{\mu\nu}] + \partial_\mu[A_\nu, F^{\mu\nu}] + [A_\mu, [A_\nu, F^{\mu\nu}]] \\ &= [A_\mu, D_\nu F^{\mu\nu}] + [\partial_\mu A_\nu, F^{\mu\nu}] + [A_\nu, \partial_\mu F^{\mu\nu}] \\ &= [A_\mu, D_\nu F^{\mu\nu}] + \frac{1}{2}[F_{\mu\nu} - [A_\mu, A_\nu], F^{\mu\nu}] + [A_\nu, D_\mu F^{\mu\nu} - [A_\mu, F^{\mu\nu}]] \\ &= [A_\mu, D_\nu F^{\mu\nu}] + [A_\nu, D_\mu F^{\mu\nu}] - \frac{1}{2}[[A_\mu, A_\nu], F^{\mu\nu}] - [A_\nu, [A_\mu, F^{\mu\nu}]] \\ &= -\frac{1}{2}[[A_\mu, A_\nu], F^{\mu\nu}] - [A_\nu, [A_\mu, F^{\mu\nu}]] \\ &= 0 \qquad\qquad (\text{ using the Jacobi identity}) \end{aligned} \tag{6.41}$$

as required. ■

6.3.1 The Yang-Mills Action

Definition 54. *The non-abelian Yang-Mills Lagrangian is*

$$\mathcal{L} = \frac{1}{4e^2}\kappa(F_{\mu\nu}, F^{\mu\nu}) \tag{6.42}$$

where κ is the Killing form of the compact matrix Lie group G.

Proposition 34. *The non-Abelian Yang-Mills Lagrangian is gauge invariant*

Proof

Under a gauge transformation

$$\kappa(F_{\mu\nu}, F^{\mu\nu}) \longrightarrow \kappa(gF_{\mu\nu}g^{-1}, gF^{\mu\nu}g^{-1}) \tag{6.43}$$

Suppose that $X, Y, Z \in \mathcal{L}(G)$. Then for $t \in \mathbb{R}$, compute

$$\frac{d}{dt}\kappa(e^{tZ}Xe^{-tZ}, e^{tZ}Ye^{-tZ}) = \kappa(-e^{tZ}[X,Z]e^{-tZ}, e^{tZ}Ye^{-tZ}) + \kappa(e^{tZ}Xe^{-tZ}, -e^{tZ}[Y,Z]e^{-tZ})$$

$$
\begin{aligned}
&= \kappa(-[e^{tZ}Xe^{-tZ}, Z], e^{tZ}Ye^{-tZ}) + \kappa(e^{tZ}Xe^{-tZ}, -[e^{tZ}Ye^{-tZ}, Z]) \\
&= -\kappa(e^{tZ}Xe^{-tZ}, [Z, e^{tZ}Ye^{-tZ}]) + \kappa(e^{tZ}Xe^{-tZ}, -[e^{tZ}Ye^{-tZ}, Z]) \\
&= 0
\end{aligned} \tag{6.44}
$$

where we have used the associativity of the Killing form. Therefore

$$
\kappa(e^{Z}Xe^{-Z}, e^{Z}Ye^{-Z}) = \kappa(X, Y) \tag{6.45}
$$

and hence it follows that

$$
\kappa(gF_{\mu\nu}g^{-1}, gF^{\mu\nu}g^{-1}) = \kappa(F_{\mu\nu}, F^{\mu\nu}) \tag{6.46}
$$

as required. ■

Note that the coupling constant e plays an important role in the dynamics. If one attempts to rescale A so that $A_\mu = e\hat{A}_\mu$, it is possible to eliminate the explicit factor of e from the Yang-Mills Lagrangian, and write

$$
\frac{1}{4e^2}\kappa(F_{\mu\nu}, F^{\mu\nu}) = \frac{1}{4}\kappa(\hat{F}_{\mu\nu}, \hat{F}^{\mu\nu}) \tag{6.47}
$$

where here $\hat{F}_{\mu\nu} = \partial_\mu \hat{A}_\nu - \partial_\nu \hat{A}_\mu + e[\hat{A}_\mu, \hat{A}_\nu]$. Although the explicit e-dependence of the Yang-Mills Lagrangian appears to have been removed, observe that the gauge field strength now has an e-dependent term, which arises from the commutator which is quadratic in A. So the dependence on e in the non-abelian theory cannot be removed by rescaling. If, nevertheless, one performs this rescaling (and then drops the ˆ on all terms), then the *generic* covariant derivative is modified via $D_\mu \Phi = \partial_\mu \Phi + ed(A_\mu)\Phi$, and the gauge potential transformation rule is also modified: $A_\mu \to A'_\mu = gA_\mu g^{-1} - e^{-1}\partial_\mu g g^{-1}$. Whether e appears as an overall factor in the Yang-Mills Lagrangian, or within the covariant derivative and gauge field strength, depends on convention. Until stated otherwise, we shall however retain the $\frac{1}{4e^2}$ outside the Lagrangian, and work with the un-rescaled gauge fields.

If the representation $\mathcal{D}$ of G is unitary, then one can couple addtional gauge invariant scalar terms to the Yang-Mills Lagrangian:

$$
\mathcal{L} = \frac{1}{4e^2}\kappa(F_{\mu\nu}, F^{\mu\nu}) + \frac{1}{2}(D_\mu \Phi)^\dagger D^\mu \Phi - V(\Phi^\dagger \Phi) \tag{6.48}
$$

where $\Phi \to \mathcal{D}(g)\Phi$ under a gauge transformation.

6.3.2 The Yang-Mills Equations

Consider a first order variation to the gauge potential $A_\mu \to A_\mu + \delta A_\mu$. Under this variation

$$
F_{\mu\nu} \to F_{\mu\nu} + (\partial_\mu \delta A_\nu + [A_\mu, \delta A_\nu]) - (\partial_\nu \delta A_\mu + [A_\nu, \delta A_\mu]) \tag{6.49}
$$

Hence, the first order variation of the Yang-Mills action $S = \frac{1}{4e^2}\int \kappa(F_{\mu\nu}, F^{\mu\nu}) d^4x$ is given by

$$
\delta S = \frac{1}{2e^2}\int \kappa(\delta F_{\mu\nu}, F^{\mu\nu}) d^4x
$$

$$= \frac{1}{e^2}\int \kappa(\partial_\mu \delta A_\nu + [A_\mu, \delta A_\nu], F^{\mu\nu}) d^4x$$
$$= -\frac{1}{e^2}\int \kappa(\delta A_\nu, \partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}]) d^4x + \text{ surface terms} \qquad (6.50)$$

where we have made use of the associativity of the Killing form. Neglecting the surface terms (assuming that the solutions are sufficiently smooth and fall off sufficiently fast at infinity), and requiring that $\delta S = 0$ for all variations δA_ν we obtain the Yang-Mills field equations

$$\partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}] = 0 \qquad (6.51)$$

or equivalently

$$D_\mu F^{\mu\nu} = 0 \qquad (6.52)$$

where here D_μ is the adjoint covariant derivative.

6.3.3 The Bianchi Identity

There is a non-abelian generalization of the identity $\epsilon^{\lambda\mu\nu\rho}\partial_\mu f_{\nu\rho} = 0$ in electromagnetism. Let D_μ be the adjoint covariant derivative. Consider

$$\begin{aligned} D_\mu F_{\nu\rho} &= \partial_\mu(\partial_\nu A_\rho - \partial_\rho A_\nu + [A_\nu, A_\rho]) + [A_\mu, \partial_\nu A_\rho - \partial_\rho A_\nu + [A_\nu, A_\rho]] \\ &= \partial_\mu(\partial_\nu A_\rho - \partial_\rho A_\nu) - ([A_\mu, \partial_\rho A_\nu] + [A_\rho, \partial_\mu A_\nu]) \\ &+ [A_\mu, \partial_\nu A_\rho] + [A_\nu, \partial_\mu A_\rho] + [A_\mu, [A_\nu, A_\rho]] \end{aligned} \qquad (6.53)$$

Consider the contraction $\epsilon^{\lambda\mu\nu\rho} D_\mu F_{\nu\rho}$. As the terms $([A_\mu, \partial_\rho A_\nu] + [A_\rho, \partial_\mu A_\nu])$ are symmetric in μ, ρ and the terms $[A_\mu, \partial_\nu A_\rho] + [A_\nu, \partial_\mu A_\rho]$ are symmetric in μ, ν, these terms give no contribution to $\epsilon^{\lambda\mu\nu\rho} D_\mu F_{\nu\rho}$. Also,

$$\epsilon^{\lambda\mu\nu\rho}[A_\mu, [A_\nu, A_\rho]] = 0 \qquad (6.54)$$

from the Jacobi identity, and just as for electromagnetism,

$$\epsilon^{\lambda\mu\nu\rho}\partial_\mu(\partial_\nu A_\rho - \partial_\rho A_\nu) = 0 \qquad (6.55)$$

because the partial derivatives commute with each other. Hence we find $\epsilon^{\lambda\mu\nu\rho} D_\mu F_{\nu\rho} = 0$, or equivalently

$$D_\mu F_{\nu\rho} + D_\nu F_{\rho\mu} + D_\rho F_{\mu\nu} = 0 \qquad (6.56)$$

This equation is called the Bianchi identity.

This can be used to prove a Jacobi identity for all covariant derivatives. Suppose that D_μ is the covariant derivative associated with the representation $\mathcal{D}$, and $\Phi \in V$ transforms under the representation $\mathcal{D}$ as $\Phi \rightarrow \Phi' = \mathcal{D}(g)\Phi$. Let d be the induced representation of $\mathcal{L}(G)$.

Then note that

$$\begin{aligned} D_\nu D_\lambda \Phi &= D_\nu(\partial_\lambda + d(A_\lambda)\Phi) \\ &= \partial_\nu(\partial_\lambda \Phi + d(A_\lambda)\Phi) + d(A_\nu)\partial_\lambda \Phi + d(A_\nu) d(A_\lambda)\Phi \end{aligned}$$

$$= \partial_\nu \partial_\lambda \Phi + d(A_\lambda)\partial_\nu \Phi + d(A_\nu)\partial_\lambda \Phi + \big(d(\partial_\nu A_\lambda) + d(A_\nu)d(A_\lambda)\big)\Phi \qquad (6.57)$$

and hence

$$[D_\nu, D_\lambda]\Phi = d(F_{\nu\lambda})\Phi \qquad (6.58)$$

Then

$$\begin{aligned} [D_\mu, [D_\nu, D_\lambda]]\Phi &= D_\mu(d(F_{\nu\lambda})\Phi) - d(F_{\nu\lambda})D_\mu\Phi \\ &= \partial_\mu(d(F_{\nu\lambda})\Phi) + d(A_\mu)d(F_{\nu\lambda})\Phi - d(F_{\nu\lambda})(\partial_\mu\Phi + d(A_\mu)\Phi) \\ &= \big(\partial_\mu d(F_{\nu\lambda}) + [d(A_\mu), d(F_{\nu\lambda})]\big)\Phi \\ &= d(D^{adj}_\mu F_{\nu\lambda})\Phi \end{aligned} \qquad (6.59)$$

where D^{adj}_μ denotes the adjoint covariant derivative.
Hence

$$\begin{aligned} \Big([D_\mu, [D_\nu, D_\lambda]] + [D_\nu, [D_\lambda, D_\mu]] + [D_\lambda, [D_\mu, D_\nu]]\Big)\Phi &= d(D^{adj}_\mu F_{\nu\lambda} + D^{adj}_\nu F_{\lambda\mu} + D^{adj}_\lambda F_{\mu\nu})\Phi \\ &= 0 \end{aligned} \qquad (6.60)$$

using the Bianchi identity on F. As this must hold for all Φ we obtain the Jacobi identity for covariant derivatives:

$$[D_\mu, [D_\nu, D_\lambda]] + [D_\nu, [D_\lambda, D_\mu]] + [D_\lambda, [D_\mu, D_\nu]] = 0 \qquad (6.61)$$

6.4 Yang-Mills Energy-Momentum

The energy momentum tensor of the Yang-Mills field strength is

$$T_{\mu\nu} = \kappa(F_{\mu\lambda}, F_\nu{}^\lambda) - \frac{1}{4}\eta_{\mu\nu}\kappa(F_{\rho\sigma}, F^{\rho\sigma}) \qquad (6.62)$$

This differs from the canonical energy-momentum tensor by a total derivative. Observe that $T_{\mu\nu}$ is gauge invariant by construction.

Proposition 35. *If the Yang-Mills field strength F satisfies the Yang-Mills field equations, the energy-momentum tensor satisfies*

$$\partial^\mu T_{\mu\nu} = 0 \qquad (6.63)$$

<u>Proof</u>

$$\begin{aligned} \partial^\mu T_{\mu\nu} &= \partial^\mu(\kappa(F_{\mu\lambda}, F_\nu{}^\lambda)) - \frac{1}{4}\partial_\nu(\kappa(F_{\rho\sigma}, F^{\rho\sigma})) \\ &= \kappa(\partial^\mu F_{\mu\lambda}, F_\nu{}^\lambda) + \kappa(F_{\mu\lambda}, \partial^\mu F_\nu{}^\lambda) - \frac{1}{2}\kappa(\partial_\nu F_{\rho\sigma}, F^{\rho\sigma}) \\ &= \kappa(D^\mu F_{\mu\lambda} - [A^\mu, F_{\mu\lambda}], F_\nu{}^\lambda) + \kappa(F_{\mu\lambda}, D^\mu F_\nu{}^\lambda - [A^\mu, F_\nu{}^\lambda]) \end{aligned}$$

$$- \frac{1}{2}\kappa(D_\nu F_{\rho\sigma} - [A_\nu, F_{\rho\sigma}], F^{\rho\sigma}) \tag{6.64}$$

where here D_μ is the adjoint covariant derivative. Note that by the Bianchi identity

$$\begin{aligned} \kappa(F_{\mu\lambda}, D^\mu F_\nu{}^\lambda) &= \kappa(F_{\mu\lambda}, -D_\nu F^{\lambda\mu} - D^\lambda F^\mu{}_\nu) \\ &= \kappa(F_{\mu\lambda}, -D_\nu F^{\lambda\mu}) - \kappa(F_{\mu\lambda}, D^\mu F_\nu{}^\lambda) \end{aligned} \tag{6.65}$$

and so

$$\kappa(F_{\mu\lambda}, D^\mu F_\nu{}^\lambda) = \frac{1}{2}\kappa(F_{\mu\lambda}, D_\nu F^{\mu\lambda}) \tag{6.66}$$

Hence

$$\begin{aligned} \partial^\mu T_{\mu\nu} &= \kappa(D^\mu F_{\mu\lambda}, F_\nu{}^\lambda) - \kappa([A^\mu, F_{\mu\lambda}], F_\nu{}^\lambda) - \kappa(F_{\mu\lambda}, [A^\mu, F_\nu{}^\lambda]) \\ &+ \frac{1}{2}\kappa([A_\nu, F_{\rho\sigma}], F^{\rho\sigma}) \end{aligned} \tag{6.67}$$

However, using the associativity of the Killing form

$$\kappa([A_\nu, F_{\rho\sigma}], F^{\rho\sigma}) = \kappa(A_\nu, [F_{\rho\sigma}, F^{\rho\sigma}]) = 0 \tag{6.68}$$

and

$$\begin{aligned} \kappa([A^\mu, F_{\mu\lambda}], F_\nu{}^\lambda) + \kappa(F_{\mu\lambda}, [A^\mu, F_\nu{}^\lambda]) &= \kappa([A^\mu, F_{\mu\lambda}], F_\nu{}^\lambda) + \kappa([F_{\mu\lambda}, A^\mu], F_\nu{}^\lambda) \\ &= 0 \end{aligned} \tag{6.69}$$

hence

$$\partial^\mu T_{\mu\nu} = \kappa(D^\mu F_{\mu\lambda}, F_\nu{}^\lambda) = 0 \tag{6.70}$$

from the Yang-Mills field equations, as required. ■

6.5 The QCD Lagrangian

The QCD Lagrangian consists of a non-abelian Yang-Mills gauge theory coupled to fermions.

The gauge group is $G = SU(3)_{colour}$. Take generators T_a of G; as G is compact, these may be chosen so that the Killing form has components

$$\kappa_{ab} = -\delta_{ab} \tag{6.71}$$

where $a = 1, \ldots, 8$ ($SU(3)$ is an 8-dimensional Lie group).

The Lagrangian is given by

$$\mathcal{L}_{QCD} = -\frac{1}{4}\sum_{a=1}^{8} F^a_{\mu\nu} F^{a\mu\nu} + i\sum_{f=1}^{6} \bar{\Psi}^A_f \gamma^\mu (\delta_A{}^B \partial_\mu + eA^a{}_\mu (T_a)_A{}^B)\Psi_{Bf} - \sum_{f=1}^{6} m_f \bar{\Psi}^A_f \Psi_{Af} \tag{6.72}$$

Here we have decomposed the non-abelian gluon field strength F and potential A into components

$$F_{\mu\nu} = F^a_{\mu\nu} T_a, \qquad A_\mu = A^a_\mu T_a \tag{6.73}$$

where

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + e[A_\mu, A_\nu]^a = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + e c_{bc}{}^a A^b_\mu A^c_\nu \tag{6.74}$$

The indices $f = 1, \ldots, 6$ are flavour indices; whereas the indices $A, B = 1, 2, 3$ are $SU(3)_{colour}$ indices. The $SU(3)_{colour}$ structure constants are $c_{ab}{}^c$. m_f is the mass of the f-flavour quark. The $SU(3)$ gauge coupling is e; observe that the gauge potentials have been re-scaled as mentioned previously in order to remove the gauge coupling factor from the Yang-Mills action; although this means that e now appears in the fundamental covariant derivative via $\partial_\mu + eA_\mu$, and also in the gauge transformations via

$$A_\mu \rightarrow g A_\mu g^{-1} - e^{-1} \partial_\mu g g^{-1} \tag{6.75}$$

where $g \in G$.

The Ψ_{Af} are fermionic fields associated to the quarks- they are 4-component Dirac spinors. We set $\Psi^A_f = \delta^{AB} \Psi_{Bf}$, and $\bar{\Psi}^A_f \equiv (\Psi^A_f)^\dagger \gamma^0$.

Gauge Invariance

The Yang-Mills term $-\frac{1}{4}\sum_{a=1}^{8} F^a_{\mu\nu} F^{a\mu\nu} = \frac{1}{4}\kappa(F_{\mu\nu}, F^{\mu\nu})$ in the QCD action is automatically gauge invariant by construction.

It remains to consider the terms involving fermions: the fermions transform under the fundamental representation of $SU(3)_{colour}$ via

$$\Psi_{Af} \rightarrow g_A{}^B(x) \Psi_{Bf} \tag{6.76}$$

where $g(x) \in SU(3)$. It is then straightforward to see that the covariant derivative $(\delta_A{}^B \partial_\mu + e(A_\mu)_A{}^B)\Psi_{Bf}$ then transforms as

$$(\delta_A{}^B \partial_\mu + e(A_\mu)_A{}^B)\Psi_{Bf} \rightarrow g_A{}^C (\delta_C{}^B \partial_\mu + e(A_\mu)_C{}^B)\Psi_{Bf} \tag{6.77}$$

Also, the $\bar{\Psi}^A_f$ transform as

$$\bar{\Psi}^A_f \rightarrow \bar{\Psi}_{Qf} \delta^{AP} (g^*)_P{}^Q \tag{6.78}$$

Hence we see that under these gauge transformations

$$\begin{aligned} \bar{\Psi}^A_f \Psi_{Af} &\rightarrow \bar{\Psi}_{Qf} \delta^{AP} (g^*)_P{}^Q g_A{}^C \Psi_{Cf} \\ &= \bar{\Psi}_{Qf} \delta^{QC} \Psi_{Cf} \\ &= \bar{\Psi}^A_f \Psi_{Af} \end{aligned} \tag{6.79}$$

where $g \in SU(3)$ implies $\delta^{AP}(g^*)_P{}^Q g_A{}^C = \delta^{QC}$.

Similarly, under this transformation

$$\bar{\Psi}^A_f (\delta_A{}^B \partial_\mu + e(A_\mu)_A{}^B)\Psi_{Bf} \rightarrow \bar{\Psi}_{Qf} \delta^{AP} (g^*)_P{}^Q g_A{}^C (\delta_C{}^B \partial_\mu + e(A_\mu)_C{}^B)\Psi_{Bf}$$

$$= \bar{\Psi}^A_f(\delta_A{}^B\partial_\mu + e(A_\mu)_A{}^B)\Psi_{Bf} \tag{6.80}$$

It follows that all terms in the Lagrangian are gauge invariant.

Equations of Motion

There are two sets of equations of motion. Varying the gauge potentials A^a_ν we obtain the variation of the action

$$\delta S = \int d^4x \; \delta A^a_\nu\big(\partial_\mu F^{a\mu\nu} + e[A_\mu, F^{\mu\nu}]^a + ie\sum_{f=1}^{6}\bar{\Psi}^A_f\gamma^\nu(T_a)_A{}^B\Psi_{Bf}\big) \tag{6.81}$$

from which we obtain

$$\partial_\mu F^{a\mu\nu} + ec_{bc}{}^a A^b_\mu F^{c\mu\nu} + ie\sum_{f=1}^{6}\bar{\Psi}^A_f\gamma^\nu(T_a)_A{}^B\Psi_{Bf} = 0 \tag{6.82}$$

Defining

$$J^\nu = -ie\sum_{a=1}^{8}\sum_{f=1}^{6}\bar{\Psi}^A_f\gamma^\nu(T_a)_A{}^B\Psi_{Bf}T_a \tag{6.83}$$

this can be rewritten as

$$D_\mu F^{\mu\nu} = J^\nu \tag{6.84}$$

Observe that $D_\mu J^\mu = 0$ as a consequence of Lemma 14.

There are also fermionic equations of motion; these may be obtained by varying $\bar{\Psi}^A_f$ and Ψ_{Bf} in the action. These can be varied independently; from the variation of $\bar{\Psi}^A_f$ (not varying Ψ_{Bf}) we obtain the equation:

$$i\gamma^\mu\big(\delta_A{}^B\partial_\mu + eA^a_\mu(T_a)_A{}^B\big)\Psi_{Bf} - m_f\Psi_{Af} = 0 \tag{6.85}$$

Next vary Ψ_{Bf} (not varying $\bar{\Psi}^A_f$)- then the action variation is

$$\begin{aligned}\delta S &= \int d^4x \left(i\bar{\Psi}^A_f\gamma^\mu\big(\delta_A{}^B\partial_\mu + eA^a_\mu(T_a)_A{}^B\big) - m_f\bar{\Psi}^B_f\right)\delta\Psi_{Bf}\\ &= \int d^4x \left(-i\partial_\mu\bar{\Psi}^A_f\gamma^\mu\delta_A{}^B + ie\bar{\Psi}^A_f\gamma^\mu A^a_\mu(T_a)_A{}^B - m_f\bar{\Psi}^B_f\right)\delta\Psi_{Bf}\end{aligned} \tag{6.86}$$

Hence we obtain

$$-i\partial_\mu\bar{\Psi}^A_f\gamma^\mu\delta_A{}^B + ie\bar{\Psi}^A_f\gamma^\mu A^a_\mu(T_a)_A{}^B - m_f\bar{\Psi}^B_f = 0 \tag{6.87}$$

Take the hermitian transpose of the above equation we obtain

$$i\gamma^0\gamma^\mu\partial_\mu\Psi^B_f - ie\gamma^0\gamma^\mu A^a_\mu(T^*_a)_A{}^B\Psi^A_f - m_f\gamma^0\Psi^B_f = 0 \tag{6.88}$$

where we have made use of the identity $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$. As T_a is anti-hermitian, we therefore recover (6.85).